UNIVERSITY OF LONDON THESIS

Degree PhD Year 2007 Name of Author JOSEPHINE BARNES.

COPYRIGHT

This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION

I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOAN

Theses may not be lent to individuals, but the University Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: The Theses Section, University of London Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION

University of London theses may not be reproduced without explicit written permission from the University of London Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

A. Before 1962. Permission granted only upon the prior written consent of the author. (The University Library will provide addresses where possible).

B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.

C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.

D. 1989 onwards. Most theses may be copied.

This thesis comes within category D.

☐ This copy has been deposited in the Library of UCL

☐ This copy has been deposited in the University of London Library, Senate House, Malet Street, London WC1E 7HU.

MRI-based measures in Alzheimer's disease and related disorders: regions of interest and automated techniques

Thesis submitted for the degree of

Doctor of Philosophy

Josephine Barnes

University College London
Institute of Neurology

2006

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

DECLARATION STATEMENT

I Josephine Barnes, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Josephine Barnes

ABSTRACT

This thesis investigates the use of manual measures of brain structure delineation on MR scans in order to assess atrophy in dementia. It further investigates the automation of atrophy measures. A new protocol for outlining the cingulate is described and was applied to groups of Alzheimer's disease (AD) and control subjects. The application of existing hippocampal and amygdala protocols to a group of pathologically-confirmed AD, frontotemporal lobar degeneration (FTLD) and controls is detailed. This analysis shows cross-sectional measurements were useful subject-group discriminators and that patterns of atrophy within and between structures may distinguish diseases.

Manual delineation of regions (cingulate and hippocampus) was extended to longitudinal studies to establish atrophy rates in groups of AD, FTLD and controls. The cingulate was shown to be at least as affected as the hippocampus by disease. Hippocampal atrophy rates from inter-scan intervals of six months were compared with measurements in the same subjects of one-year intervals. Results from studies combining cross-sectional and longitudinal hippocampal data are described. These assess the asymmetry of the structure and investigate the pre-symptomatic decrease in volume in familial AD subjects.

Semi-automated techniques were performed which utilise registration of serial hippocampi to assess change longitudinally. Results show the semi-automated techniques to be reliable and consistent with manual measures. These techniques were then applied to scans from a multi-centre clinical trial and again consistency with manual measures was assessed.

The generation of fully-automated template-based hippocampal segmentations is described. The approximate regions generated from the template were used to quantify the boundary shift integral and the resultant atrophy rates were compared with manual rates revealing automated measures to be consistent with manual measures.

These results are put within the context of other similar studies by performing a meta-analysis of hippocampal atrophy rates. Implications for diagnosis and monitoring disease progression are discussed.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

AIMS OF THIS THESIS

The Problem

Alzheimer's disease (AD) is a growing socio-economic and healthcare problem affecting approximately 20% of the world's population over 80 years of age (Dawbarn and Allen, 2001). Structural MR imaging is a non-invasive means of assessing one result of the AD pathology, namely atrophy. Structures of the brain such as the hippocampus are affected early in the disease (Braak and Braak, 1991), and as a result there is great interest in measuring this structure at one time-point or using multiple time-points, to assess change. There is also evidence that another brain structure, the cingulate gyrus, is also affected in AD (Braak *et al.*, 1993). Measurement of such substructures may aid: a) differentiation of dementia from normally ageing subjects, b) the differential diagnosis of dementia, c) the ability to track disease over time, and d) our understanding of the natural history of the disease. Manual measures of such structures are the "gold standard" for assessing volume at one time-point or change in volume over time. This type of analysis is both time-consuming and requires highly-trained operators and as a result this assessment is restricted to specialist centres. The increased automation of such measures may improve the utility of this type of volumetric analysis in a clinical situation and may prove useful in clinical trials of potentially disease-modifying therapies.

Aims

1. To assess the existing manual delineation protocols of the hippocampus cross-sectionally and longitudinally in dementia in terms of differential diagnosis and distinguishing AD subjects from controls.
2. To develop and assess manual segmentation protocols of the cingulate gyrus in order to analyse volumes at one time-point and change over time.
3. To assess semi-automated measures of hippocampal atrophy rates in their ability to separate AD subjects from controls. These measures are to be compared with the "gold standard" manually-derived rates.
4. To develop and assess automated means of measuring hippocampal change over time which may have utility for clinical situations or clinical trials. Again these automated measures are to be compared with the "gold standard" manually-derived rates.
5. To assess location and distribution of atrophy within and between brain structures in order to aid differentiation of diseases and understand their natural history.

1. IMAGING CEREBRAL ATROPHY IN DEMENTIA

1.1. Introduction

Alzheimer's disease (AD) is the most common cause of dementia affecting over 5% of the population above the age of 60 (Dawbarn and Allen, 2001). The prevalence of AD doubles every five to ten years above the age of 60 (Small *et al.*, 1997). In 2000, 15% of the population in Europe was aged over 65 years and nearly 3% over 80 years. By 2030 it is projected that these numbers will increase to 24% and 6% respectively (Kinsella and Phillips, 2005). With this ageing population, AD poses a serious and growing public health problem and increasing economic burden. For each affected individual and their carers and families AD represents a devastating personal loss.

With symptomatic treatments for AD available and development of disease modifying therapies underway, accurate diagnosis of patients with AD as well as methods for monitoring disease progression are essential to allow maximum benefit to be derived from treatment.

At the earliest stages of AD, cognitive symptoms as well as brain appearances on neuroimaging often overlap with those seen in the normally-ageing population. An additional challenge relates to differentiation of changes seen in AD from those of other dementias. For a minority of AD patients a gene has been identified which allows early and even pre-symptomatic diagnosis. However in the majority, those with sporadic AD, a definite diagnosis is only possible with visualization of the characteristic histopathological appearances of the disease in cerebral tissue usually at post-mortem.

The assessment of cerebral atrophy is increasingly seen as valuable in identifying a pattern of atrophy that helps distinguish AD from other causes of dementia and normal ageing. Rates of atrophy provide a marker of, and in some cases predict, disease progression. This chapter focuses on how these advances in imaging techniques and applications have been made and assesses their current and potential applications.

1.2. Dementia

Dementia can be defined as an acquired deficit in multiple domains of cognitive function, including memory in the presence of normal consciousness (Rossor, 1993). This broad clinical syndrome may be generated by a large number of causes including neurodegenerative diseases such as AD.

1.2.1. Alzheimer's Disease

1.2.1.1.Pathology of AD

AD is characterised by progressive accumulation of extra-cellular aggregated ß amyloid (Aß) plaques and intra-cellular neurofibrillary tangles (NFT) made of abnormally hyperphosphorylated tau protein. This pathology typically begins in the entorhinal cortex and progresses to involve the hippocampus before affecting the cortex widely (Braak and Braak, 1991). As a result of these pathological processes, biological changes occur in the brain on a microscopic level including neurotransmitter loss, dendritic pruning of neurons and neuronal cell death. One consequence of these changes at a macroscopic level is cerebral atrophy.

1.2.1.2.Clinical Presentation of AD

Alzheimer's disease most commonly presents with impairment of episodic memory which progresses to involve other cognitive domains. A "definite" diagnosis of AD can only be given when brain tissue is examined pathologically; until such time, clinical diagnosis can only be termed either "possible" or "probable" according to NINCDS-ADRDA criteria (McKhann *et al.*, 1984).

Alzheimer's disease can be divided into sporadic (SAD) and familial (FAD) forms. The vast majority of cases is sporadic and has a late onset (over 65 years of age). A smaller proportion of the disease is young onset (before 65 years of age) and sporadic. An even smaller proportion is young onset and classed as familial. In this sense, familial refers to a disease with autosomal dominant inheritance usually with a predictable age of disease onset. Mutations in three genes have been found to be causative for autosomal dominant AD: these are amyloid precursor protein (APP) and presenilin 1 and 2 (PS1 and PS2). These young onset familial cases account for much less than 5% of all cases.

1.2.2. *Mild Cognitive Impairment*

Mild cognitive impairment (MCI) represents a transitional stage between normal ageing and dementia. A diagnosis of MCI is made where a person has cognitive complaints and evidence of cognitive deficits, which are insufficient to fulfil criteria for dementia. MCI can be classified as amnestic, multiple cognitive domains, single non-memory cognitive domain or vascular. The term amnestic mild cognitive impairment (*a*MCI) has been proposed to describe the major sub-group of MCI where memory deficits predominate in the absence of global cognitive and functional impairment. Subjects with *a*MCI have been reported to have an annual conversion rate to a formal diagnosis of AD of 6 to 25% (typically around 10%) compared with 0.2 to 4% in the age-equivalent general population (Petersen *et al.*, 2001a). This disease group is important as the people within it are potentially at an early stage of the neurodegenerative disease process. As such they may derive most benefit from a disease modifying therapy for AD, should it become available.

1.2.3. *Frontotemporal lobar degeneration (FTLD)*

A large number of non-AD pathological processes can cause dementia including dementia with Lewy bodies (DLB) vascular disease (VaD), Creutzfeld Jakob disease, human immunodeficiency virus, corticobasal degeneration, and progressive supranuclear palsy amongst others. Frontotemporal lobar degeneration (FTLD) describes a group of neurodegerative conditions which cause atrophy predominantly in the frontal and anterior temporal cortex. Distinction between AD and FTLD is important, as the treatments and prognoses can be very different. There are a number of different clinical syndromes associated with FTLD and numerous histological and immunohistochemical classifications of underlying pathology. The clinical presentation does not necessarily reflect the pathology; however there is a relationship between the distribution of atrophy and the clinical syndrome. For a comprehensive review of FTLD, see Neary *et al.* (Neary *et al.*, 2005).

1.2.3.1.FTLD Pathology

FTLD can be subdivided into a number of disease types (see Figure 1.1). These can be described by the following histological and immunohistochemical features on inspection:

A) Dementia lacking distinctive histological features (DLDH), with microvacuolation of the upper cortical layers present on histological inspection but no other pathological features.

B) Frontotemporal lobar degeneration of the ubiquitinated type (FTLD-U) with presence of microvacuolation in the upper cortical layers and with rounded ubiquitin-reactive only neurounal inclusions and dystrophic neurites within the frontal or temporal neocortex or the hippocampal dentate cells.

C) Pick type disease (PiD) with tau-reactive rounded intraneuronal inclusions (Pick bodies) and swollen achromatic neurons (Pick's cells).

D) Mutations in the tau gene on chromosome 17 (FTDP-17) causing microvacuolation and tau reactive NFTs and / or Pick-like bodies, and, on occasion, tangles in the glial cells in the cortical white matter.

1.2.3.2.Clinical Presentations of FTLD

Frontotemporal dementia or frontal variant frontotemporal dementia (fvFTD) is associated with a progressive behavioural syndrome which is associated primarily with atrophy of the frontal and anterior temporal lobes. Some of these frontal patients also develop motor-neuron disease (MND) and these cases are often named FTD-MND. Clinical criteria for discriminating these cases from AD were developed by the Lund and Manchester groups (The Lund and Manchester Groups, 1994). These criteria do not include other subtypes of FTLD. Progressive non-fluent aphasia (PNFA) is a term used to describe those patients with progressive speech production problems with preserved word comprehension. This is associated with atrophy focussed particularly on the left perisylvian region (Mesulam, 2001). Semantic dementia (SD) is associated with progressive problems with comprehension of words and naming of objects and other sensory stimuli. This is associated with severe and often asymmetric (left > right) atrophy of the temporal lobes. Neary (Neary *et al.*, 1998) recognised these syndromes of FTLD and these were included in guidelines published by McKhann *et al.* (Mckhann *et al.*, 2001). There is one further syndrome which may be classed a specific clinical subtype of FTLD which is not specifically delineated in existing criteria; focal right temporal lobe atrophy associated with behavioural changes, and topographical disorientation combined with prosopagnosia (Evans *et al.*, 1995; Rosen *et al.*, 2006; Thompson *et al.*, 2004c).

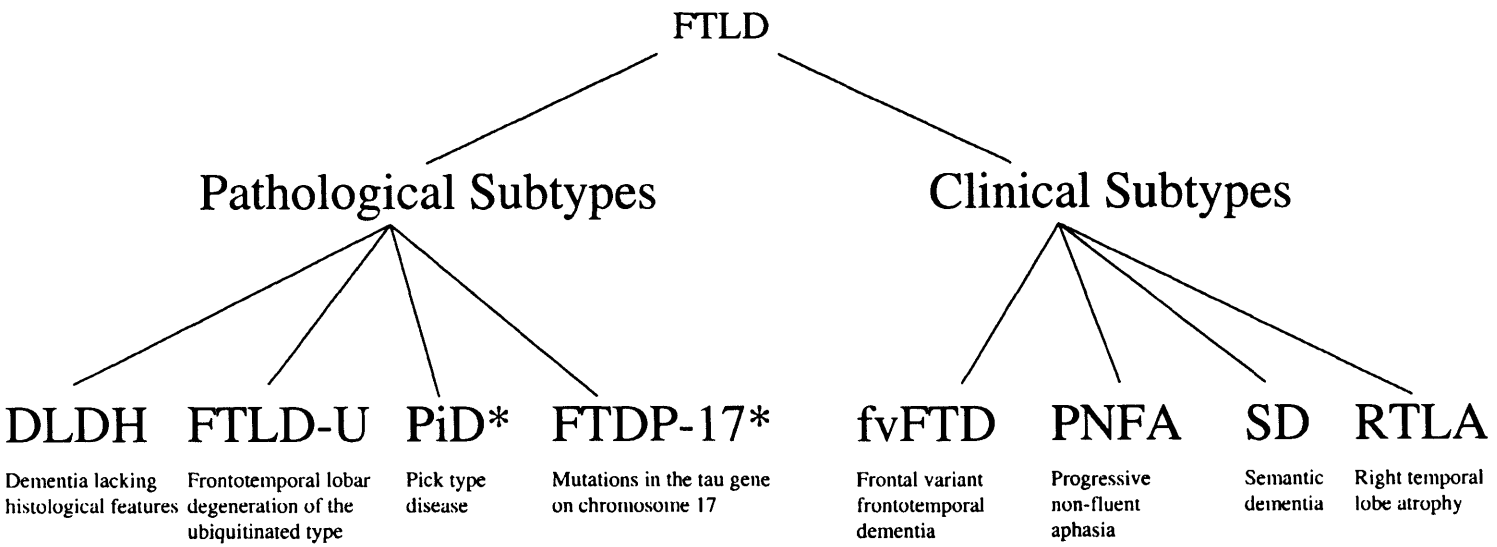


Figure 1.1 Pathological and clinical subtypes of FTLD.
**These pathological subtypes are often referred to as tauopathies.*

1.3. Imaging techniques

With the development of different imaging techniques biological features can be assessed (see Figure 1.2). Brain atrophy can be visualised using structural neuroimaging, and what this reveals about AD will be the focus of this chapter. However, for completeness a brief description of other imaging techniques which reveal additional biological changes or deficits in the AD brain will be included.

1.3.1. Functional Imaging Techniques

A number of different techniques come under the umbrella of functional imaging. These techniques include single photon emission computerised tomography (SPECT), positron emission tomography (PET), proton magnetic resonance spectroscopy (MRS), diffusion tensor imaging, and functional magnetic resonance imaging (fMRI). Each technique is applied to reveal different information about the functioning of the brain (see Figure 1.2). One method utilises PET which measures metabolism through quantifying uptake of labeled glucose (^{18}FDG-PET) over the whole brain (Alexander *et al.*, 2002). Patterns of ^{18}FDG-PET uptake can aid the differentiation of dementias: decreased temporoparietal uptake of ^{18}FDG is found in AD, whereas in the frontotemporal lobar dementias there is decreased ^{18}FDG uptake in the frontal, anterior temporal and medial temporal cortices. SPECT has shown similar results to ^{18}FDG-PET in AD and a comparison of the two methods is reviewed in (Silverman, 2004). One of the most fast-moving research areas is molecular imaging where radiotracers such as the 'Pittsburgh-B compound', ^{11}C-PIB has been used in PET studies to visualise amyloid *in vivo* differentiating patients with mild AD from controls (Klunk *et al.*, 2004). A growing area of research investigates how to combine information gained by different modes of imaging to give a full picture of the functioning and structure of the brain and its sub-regions (for a review see (Galanaud *et al.*, 2003)). As this thesis is concerned with structural MR imaging, no more detail about these functional or multi-modal techniques are given, however for overall reviews of imaging techniques used in dementia, see (Masdeu *et al.*, 2005; Petrella *et al.*, 2003).

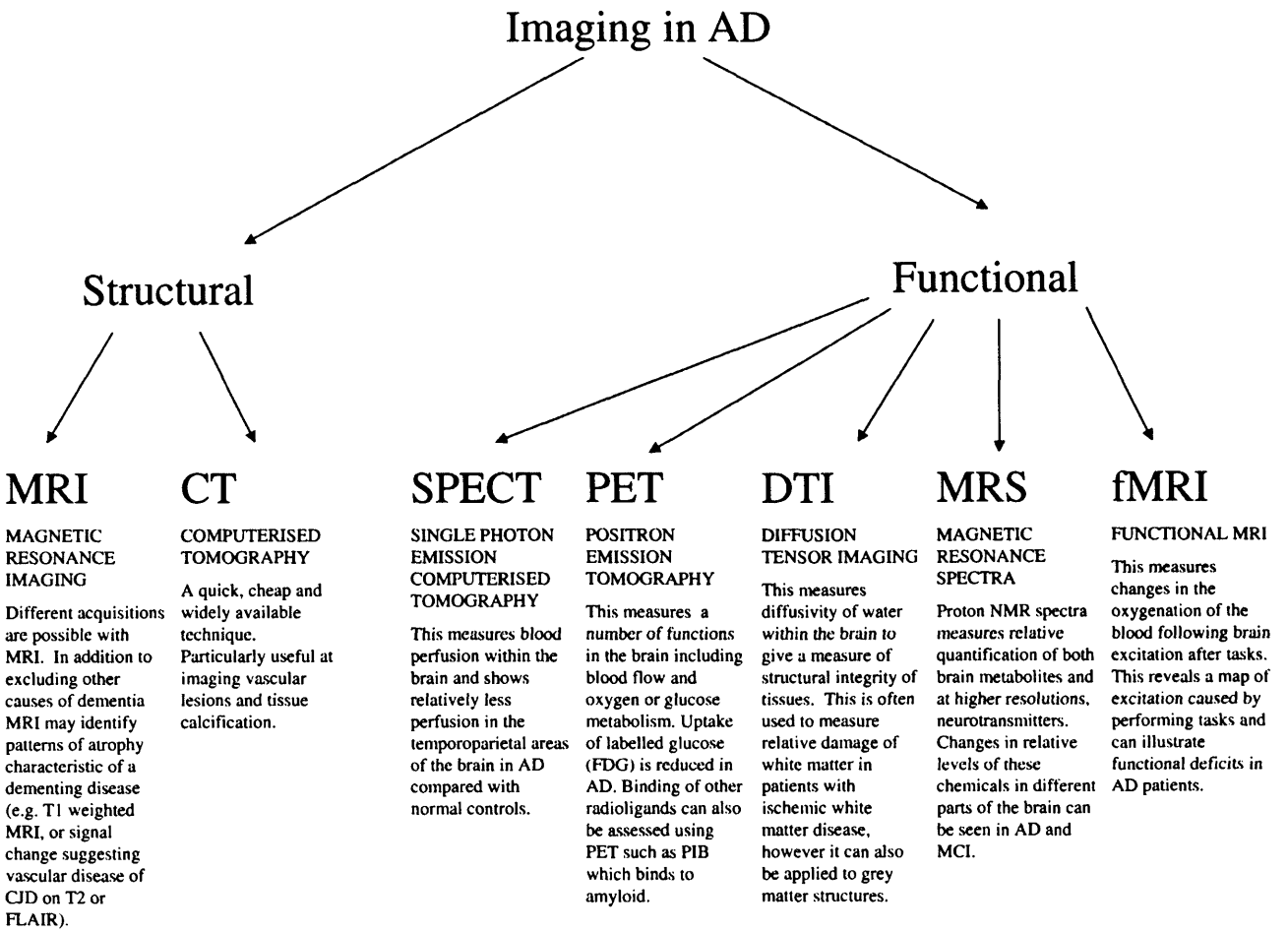


Figure 1.2 Overview of the different imaging methods used to assess dementias.

1.3.2. Structural Imaging Techniques

Structural imaging allows the visualisation of brain anatomy. Either magnetic resonance imaging (MRI) or computed tomography (CT) can be used for this type of investigation. Choice of imaging modality is often dependent on many factors. CT is cheap and widely available. The short acquisition times required by CT may be advantageous when scanning patients who have difficulty keeping still. However, the radiation dose required to acquire an image is relatively high and currently the image resolution and tissue contrast is inferior to MR. CT scanning technology has become increasingly sophisticated in recent years however MR is still considered to be the modality of choice for imaging atrophy. MR has been established as a safe, non-invasive and high resolution means of brain imaging and is increasingly available in clinical practice. MR avoids radiation but scanners are not so widely available and the images are sensitive to patient movement. In addition, the powerful magnetic fields used make this an unsuitable investigation for those with pacemakers or metal implants. European and US guidelines for the investigation of possible dementia recommend that all patients undergo either MRI or CT if possible (Knopman *et al.*, 2001; Waldemar *et al.*, 2000). Neuroimaging is important as it excludes rare but potentially treatable causes of cognitive problems, such as brain neoplasms, subdural haematomas, and aids in the diagnosis of vascular and degenerative causes of dementia (Scheltens *et al.*, 2002).

1.3.2.1. CT and MRI

CT and MR imaging are acquired in very different ways and can give differing information about the structure imaged. CT uses X rays to create an image that reflects absolute tissue density. With MR, tissue types can be differentiated by their relatively different T1 and T2 values. Cerebrospinal fluid (CSF) appears dark on T1 images but bright in proton density and T2 images. See Figure 1.3 which shows a T1-weighted MR image and a T2-weighted MR image.

As in functional imaging, high resolution structural MRI scans have been used to image amyloid plaques. Plaques have been observed in both mice (Jack, Jr. *et al.*, 2004) and human post-mortem brain samples (Benveniste *et al.*, 1999). However, this high resolution imaging requires long scanning times together with high field strengths which can not be used *in vivo*. Consequently, current MRI cannot be used to image the histopathological features of AD in humans *in vivo* because even large amyloid plaques

(approx. 100 microns in diameter) are ten times smaller than the typical resolution (approx 1mm) of clinical imaging. Nevertheless, MR imaging allows visualization of cerebral atrophy, the macroscopic effect of the pathological process in AD. The degree of cerebral atrophy measured on MRI over time has been found to correlate with the pathology found at post mortem (Silbert *et al.*, 2003).

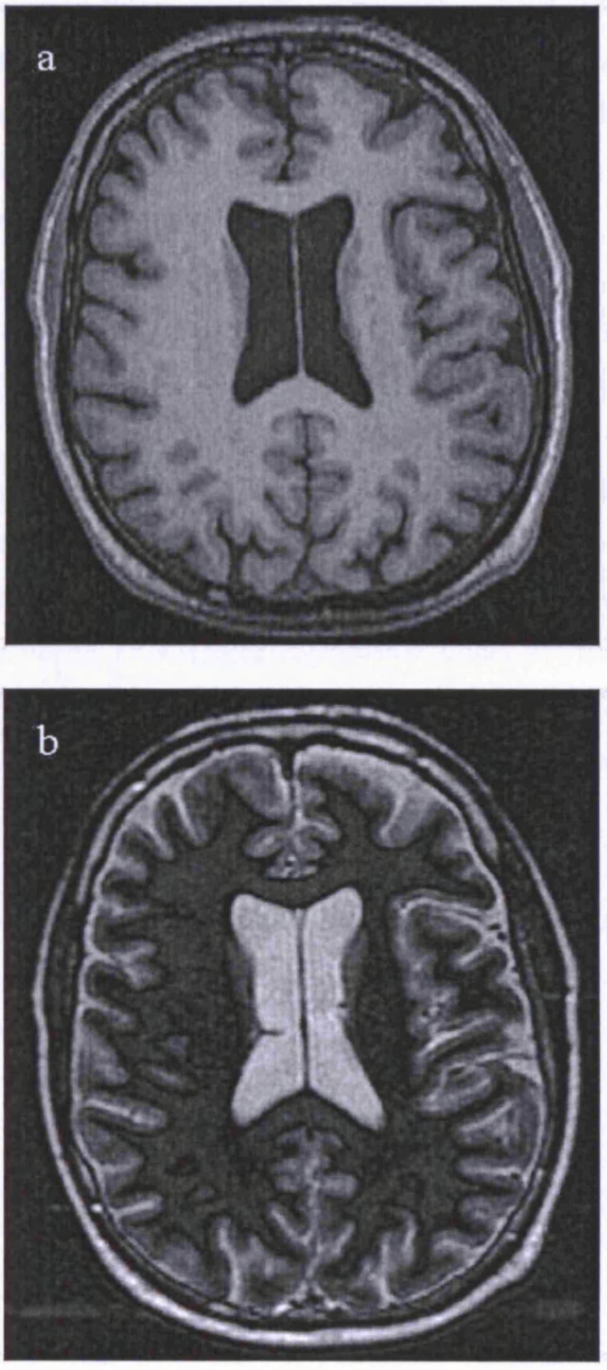


Figure 1.3 a) A "slice" of a T1-weighted MR image and b) T2-weighted MR image of the brain.

1.3.3. Cross-sectional imaging techniques

1.3.3.1. Visual inspection

The most common form of analysis in clinical practice for both MRI and CT images is simple visual inspection. Different diseases can present with different patterns of atrophy which can be visualised on CT and MR. AD is characterised by bilateral small hippocampi in the earlier stages followed by generalised cerebral atrophy, whereas FTLD presents with a more asymmetric pattern with more frontal and temporal atrophy and less posterior atrophy, see Figure 1.4. The gradient of atrophy in these diseases may help differentiate them from each other (more posterior loss in AD compared with more anterior loss in FTLD). However, patterns of atrophy do not always differentiate one disease from another, for example AD and DLB overlap in their appearances as both show generalised posterior atrophy; although DLB shows relatively less hippocampal atrophy.

The characteristic hallmarks of some diseases may be easier to identify with different modalities, for example the vascular lesions present in vascular dementia are better demonstrated on T2-weighted MR images including FLAIR (fluid attenuated inversion recovery) rather than using T1-weighted acquisition. Although visual inspection is the most common form of scan analysis, and can successfully identify certain types of pathology, the sensitivity and specificity for certain diagnoses can be low. This is often due to co-existence of pathologies (e.g. AD and vascular dementia) or the overlapping appearances of differing diseases (e.g. AD and DLB and with normal ageing). The great inter-individual variability seen in brain morphology means that the subtle patterns of atrophy seen in the earliest stages of a disease may be hard to detect. It may be only when the disease is more advanced that a characteristic pattern of atrophy can be visualised, which is suggestive of a particular diagnosis.

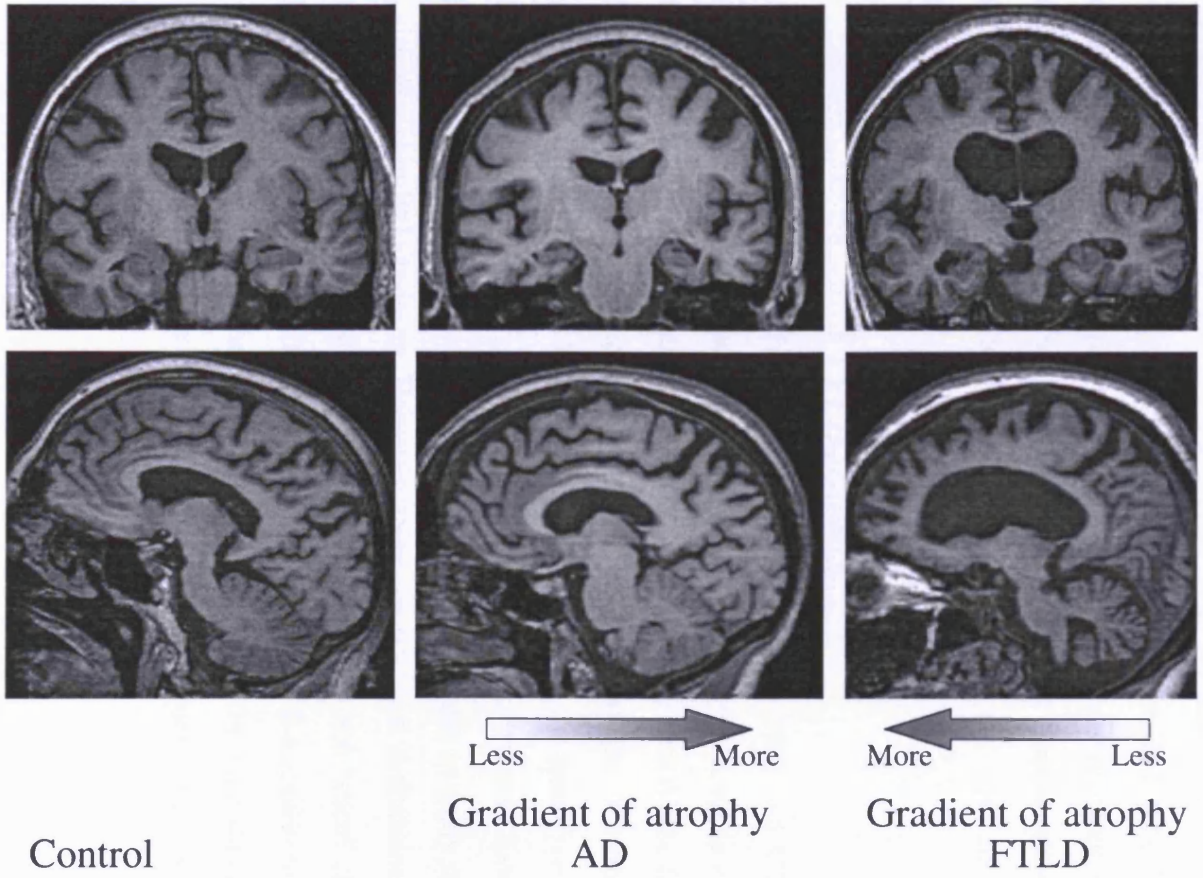


Figure 1.4 Cross-sectional MR images showing a normal control, an AD subject and an FTLD subject.
AD is characterised by relatively small but symmetrical hippocampi. FTLD can present with either frontal or asymmetric temporal atrophy, the example shown is of an fvFTD subject. Differences between disease patterns are confounded by the large degree of inter-individual variability, for example this normal control has relatively large asymmetric ventricles.

To quantify the characteristic differences seen with neuroimaging, specific visual rating scales for different markers of disease have been developed. In AD the reduction in the size of the medial temporal lobe can be semi-quantitatively estimated using the Scheltens scale (Scheltens *et al.*, 1992) (see Figure 1.5). This technique grades the scans relative to a template scale according to size of the choroidal fissure, height of the hippocampus, size of the temporal horn and collateral sulcus. This method is reproducible both between and within raters (Scheltens *et al.*, 1992). For VaD a number of scales have been developed to quantify the ischemic vascular load seen on MR scans (Scheltens *et al.*, 1998). These methods are useful in the clinical setting as they are easy to apply by a rater, require minimal training before use, and have been validated extensively.

1.3.3.2.Linear measures

Simple quantification of the cerebral atrophy seen on MRI and CT imaging can be achieved using linear measures of brain regions. Techniques applicable to CT include measurement of the minimum thickness of the medial temporal lobe (Jobst *et al.*, 1992) or the radial width of the temporal horn (Frisoni *et al.*, 2002b). In terms of MR imaging, techniques including the width of the temporal horn of the lateral ventricles have been described (Frisoni *et al.*, 1996c). Some of these techniques have been validated extensively and much like visual ratings are quick and easy to apply in a clinical setting with relatively little training. However, there are potential limitations with this type of technique including inconsistencies and dependence on the "slice" chosen which may lead to relatively poor sensitivity and specificity (e.g. 79% and 69% respectively for AD vs. normal ageing and other dementias for height of the hippocampus (Pucci *et al.*, 1998), 75% and 93% respectively for AD vs. normal controls for width of the medial temporal horn (Frisoni *et al.*, 2002b)).

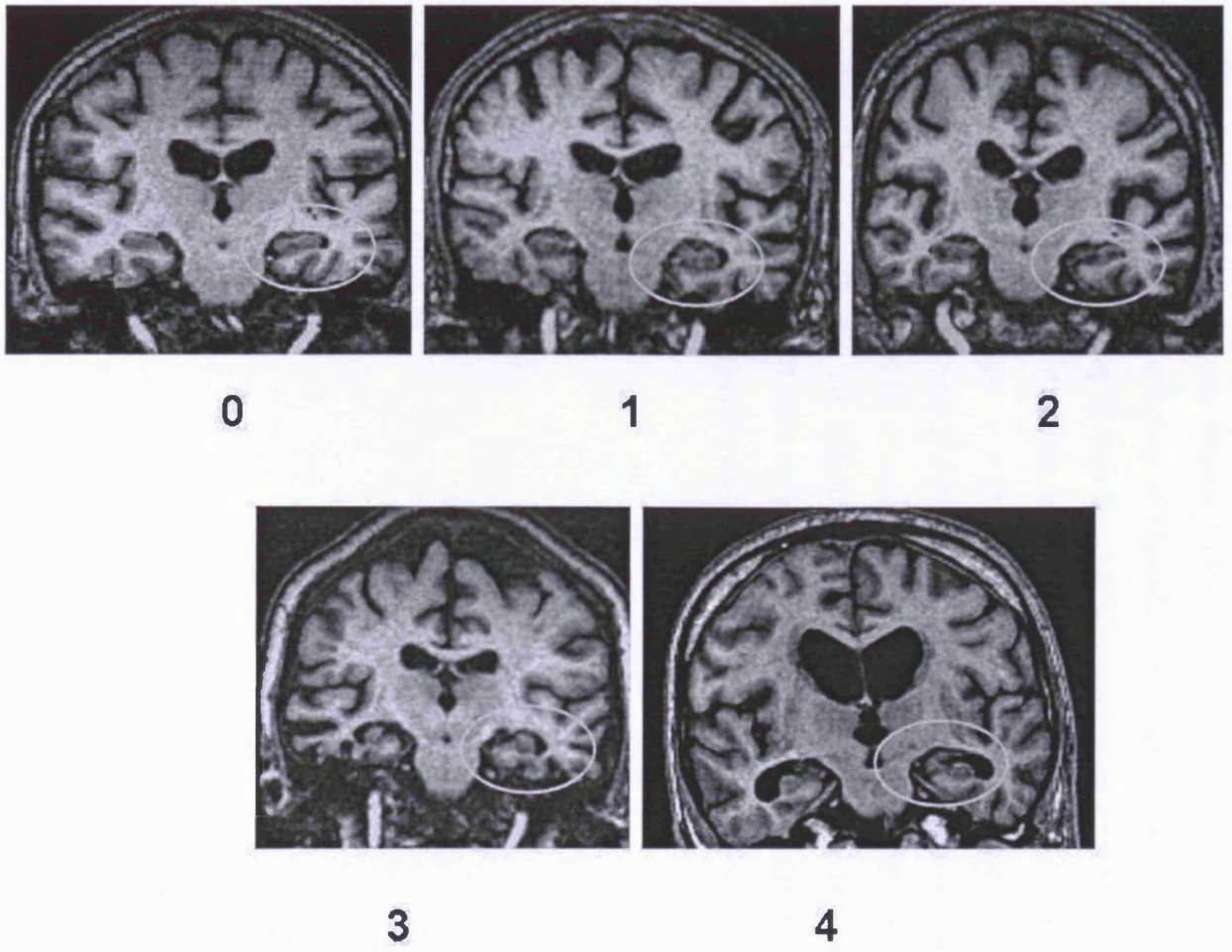
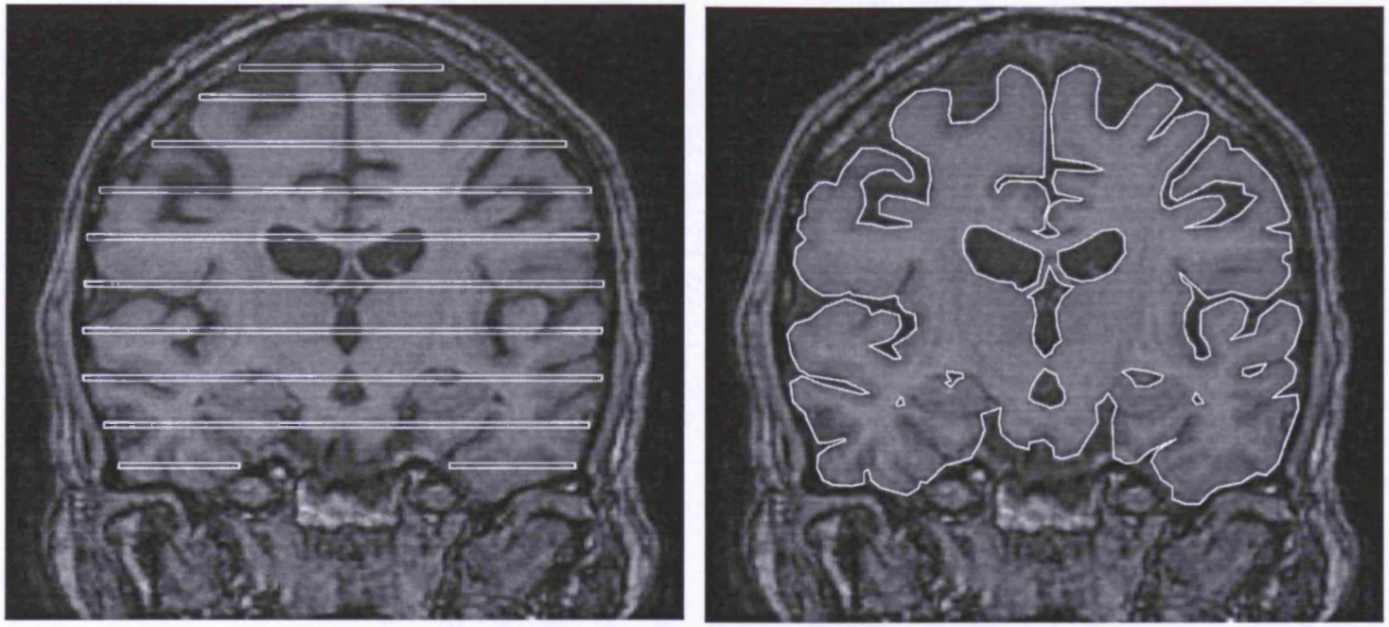


Figure 1.5 Scheltens scale of medial temporal lobe atrophy.
This scale assesses characteristics of temporal lobe atrophy within the area circled. Rating involves evaluation of the height of the hippocampus, and the widening of the choroidal fissure, temporal horn and collateral sulcus (scale 0-4, with 0 being normal and 4 being atrophied).

1.3.3.3. Volume measures

The remainder of this chapter will focus on MR imaging. Cross-sectional comparisons of brain volume give approximate insight into the amount of brain atrophy that may have occurred during the disease process. Analysis may be performed on a computer workstation using software which enables the user to outline structures in an automated or semi-automated fashion (see Figure 1.6) (Freeborough *et al.*, 1997; Smith *et al.*, 2001). Although measurements may be reproducible between research groups there is often a large degree of overlap in brain volume between AD patients and controls. This overlap may be reduced with correction for estimated pre-morbid brain size, using measures such as total intracranial volume (TIV) (Whitwell *et al.*, 2001). This type of correction does not remove overlap altogether due to the large degree of inter-individual variability in volume of brain structures. Whole brain volume analysis gives a measure of general brain atrophy but no indication of how one region of the brain has changed compared to another. As different dementias are characterised by differing patterns of atrophy, this measure is not specific enough to be diagnostically useful.



TIV Brain

Figure 1.6 Segmentation of whole brain and total intracranial volume (TIV).
Whole brain measurement is often performed on every "slice" of an MR scan. This gives an accurate estimate of brain volume which can be highly reproducible between raters (less than 1% mean absolute difference). TIV is a more regular shape and therefore may be estimated using a proportion of the slices or thicker slices. As a result, TIV is often measured at regular intervals of the scan, for example, every 10 slices. This gives an estimate of premorbid brain size / head size which can be used to correct other volumetric measures, removing inter-individual variability such as sex differences. TIV is also highly reproducible between raters (less than 1% mean absolute difference).

1.3.3.4.Regional measures of atrophy

A wealth of regional volumetric studies dating back to the early 1990s involved manually outlined smaller brain structures or regions of interest (ROI) which quantified the patterns of regional atrophy seen on the MR images. Figure 1.7 demonstrates segmentation of the hippocampus and amygdala. These structures have been of interest for both early diagnosis and progression of disease. Similar to whole brain measures, TIV correction of these volumes reduces variability between individuals and increases separation between normal controls and AD groups.

1.3.3.5.Computational anatomy: comparing anatomy at the group level

In recent years the field of computational anatomy has developed, resulting in automated and semi-automated techniques that can assess brain volumes and shapes both within and between subject groups. This can be achieved on the scale of the whole brain, or on smaller structures within the brain.

1.3.3.6.Voxel Based Morphometry (VBM)

Voxel based morphometry (VBM) is a statistically-driven image analysis package which is widely used to analyse large numbers of scans on a group level. VBM provides an environment in which scans of differing subject groups can be compared on a voxel by voxel basis throughout the image. In this way no *a priori* assumptions are made about which areas of the scan may be affected by a disease. Such analysis requires many pre-processing steps including spatial normalisation, segmentation, modulation and smoothing which results in production of tissue density maps for each group. For a detailed description of these methods see Ashburner *et al.* (Ashburner and Friston, 2000). Using these methods tissue density maps can be compared using t tests at every voxel with pre-selected significance cut-offs and multiple comparison correction. These methods have allowed comparison of patterns of atrophy and tissue loss between AD and normal ageing (Baron *et al.*, 2001; Karas *et al.*, 2003; Rombouts *et al.*, 2000). Differences in group sizes, disease severity, image quality, arbitrary significance cut-offs, and multiple comparison correction techniques have led to some discrepancies in findings between research groups.

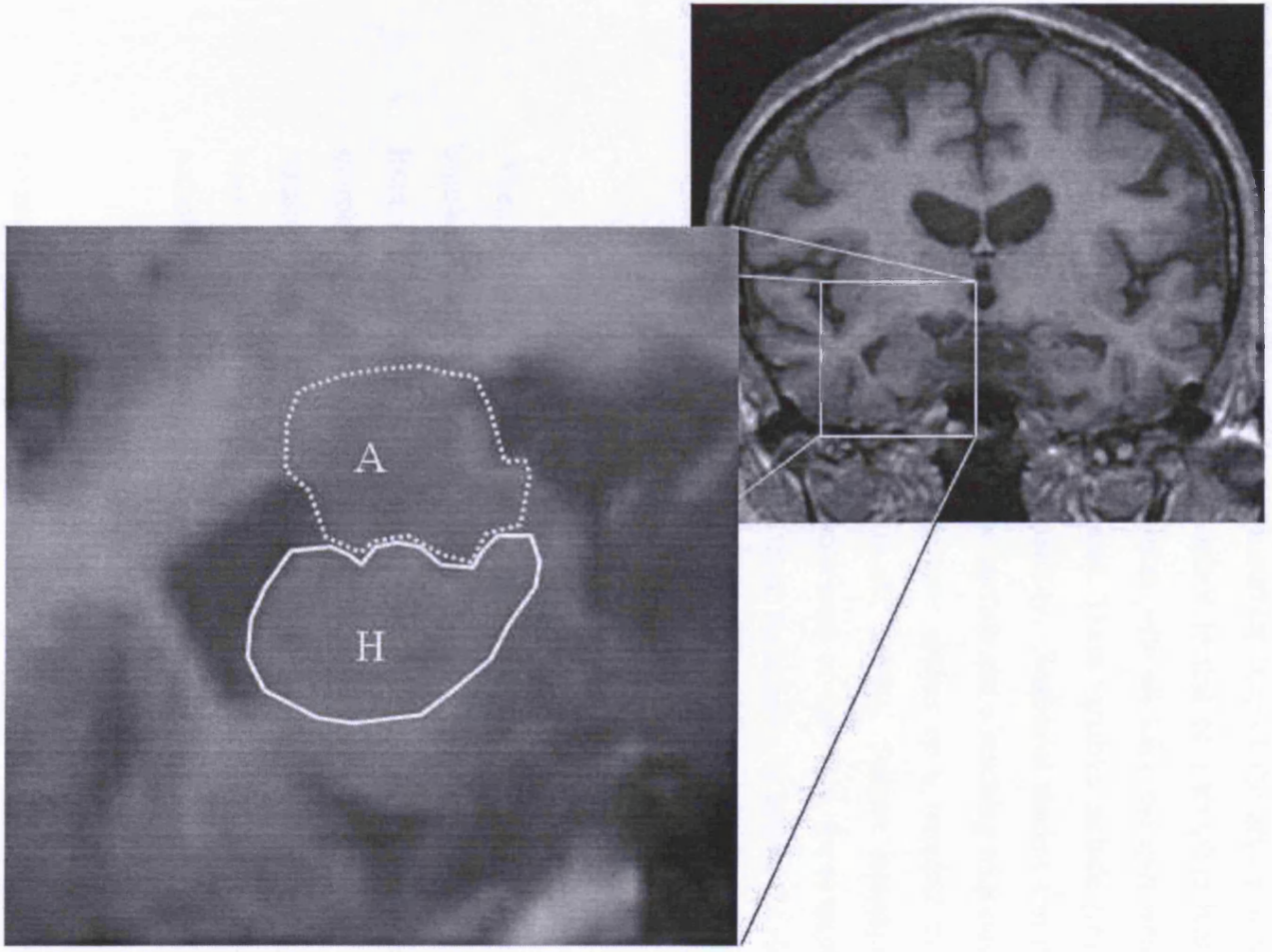


Figure 1.7 Segmentation of hippocampus (H) and amygdala (A).
These segmentations are usually performed on every "slice" the structure is present in an automated or semi-automated fashion (using pre-set intensity thresholds determined by mean whole-brain intensity). These areas are known to be affected early in the disease over and above whole brain atrophy, however the error in their measurement is relatively high (approximately 5% of the structure volume.)

1.3.3.7.Surface-based techniques

Cortical mapping is another technique which assesses the whole brain and involves several pre-processing steps including: segmentation of the brain into grey matter, white matter, and CSF, creation of a surface map of the grey matter-CSF border, and non-linear transformation of the surface to that of a template brain. Once all brains are within the same spatial framework with all sulci and gyri well matched, a number of different variables can be assessed. These variables include grey matter density, cortical thickness or gyral pattern variability. Statistical models can then be applied to these variables at all the points of the surface and a resulting map can display that variable or differences in the variable between groups on a template brain. For a review see Thompson *et al.* (Thompson *et al.*, 2004b). Surface techniques can also be used to assess differences in smaller structures of the brain, for example the hippocampus. A number of studies have used these techniques to assess the differences in shape and volume of hippocampi between AD and control groups (Csernansky *et al.*, 2000).

1.3.4. Longitudinal Measures

1.3.4.1.Introduction

When comparing scans across different subjects there is inherently a large degree of inter-individual variability. Comparison of two scans taken at different time-points from the same individual reduces the problem of the between-subject variability in morphology. In this case each patient acts as their own control and information can be obtained regarding disease progression. Quantification of change allows calculation of rates of cerebral atrophy, and improves differentiation between normal ageing and pathological processes (Fox *et al.*, 1996a; Jack *et al.*, 1998).

1.3.4.2.Registration and measuring change in the whole brain

To improve the visual assessment of serial images, scans can be co-registered (spatially matched) using a variety of techniques with differing levels of complexity. The registration process utilises software to enable one scan to be matched to the other. As well as allowing better visual assessment of obvious changes, more subtle changes may become apparent. Registration allows better localisation of atrophy than visual assessment alone and thus allows more precise measures of subtle changes. This matching also allows whole brain changes to be calculated semi-automatically, by co-registering the images and then estimating the change between the two brain regions

using a subtraction image (baseline scan – registered-repeat scan). These registration-based techniques have been used extensively to measure change in the whole brain. The technical details regarding registration and longitudinal brain change quantification are described in Chapter 2.

1.3.4.3.Measuring change in focal areas – manual measurement

In addition to measuring whole-brain volume change over time, it is possible to measure smaller areas within the brain. Changes in the CSF spaces such as the lateral ventricles provide markers of brain change over time and have been used in clinical trials owing to the ease of segmentation and reproducible nature of protocols employed. Although these features on MR are sensitive markers of disease progression, they are not specific to any particular disease, as change in ventricle size does not localise the loss of brain tissue to specific regions.

Focal changes within the brain can be detected longitudinally using manual segmentation (Du *et al.*, 2004; Jack *et al.*, 1998; Kaye *et al.*, 1997). Levels of change in these areas may also be higher than whole brain change. However, many of these measurements have an inherently higher level of measurement error compared with whole brain. This is due to difficulty in reproducible anatomical localisation and delineation of these structures, and standardization of segmentation protocols.

1.3.4.4.Group analyses: Longitudinal VBM and cortical mapping

Packages such as VBM, mentioned earlier in the text, are also able to perform group analyses of change over time. This can be achieved either by comparing the same group at baseline and assessing significant changes at repeat imaging (essentially looking cross-sectionally at both time-points); or by importing the deformations from non-linear registrations into VBM and assessing statistical differences between groups (Scahill *et al.*, 2002). An example of the latter analysis is demonstrated below in Figure 1.10. Cortical mapping of grey matter density (described earlier) has also been applied longitudinally in a group of AD revealing similar results to longitudinal VBM in AD (Thompson *et al.*, 2003).

1.4. Clinical applications of techniques – imaging a continuum

This section discusses how research using many of the techniques described above has shed light on the progression of AD from pre-symptomatic stages of the disease, to amnestic MCI, and finally conversion to AD. Figure 1.8 illustrates this spectrum of clinical progression and papers which have demonstrated features of the disease at each point.

1.4.1. Normal ageing

Cerebral structures change over time as a result of apparently healthy ageing. It is unclear how much of this is due to an inevitable intrinsic ageing process as opposed to low levels of damage due to multiple pathologies. In order to discriminate "normal" age-related changes from those associated with neurodegenerative disease, it must be known in what way this normal change occurs. However there is often difficulty in defining a cohort that reflects the 'normal population'. Study cohorts vary regarding cognitive ability, vascular risk factors, and 'normal' scan appearances. Initial studies of normal ageing involved qualitative comparison of MRI scans through visual assessment. These studies aimed to distinguish between subjects of differing age based on the prominence of certain characteristics including sulcal depth, signal hyperintensities and ventricular size. With the development of quantitative imaging techniques cross-sectional imaging studies revealed differences in the brain in elderly populations compared with that of younger populations. These changes included smaller whole brain, temporal lobe and hippocampal volumes, together with greater ventricular volumes (Gur *et al.*, 1991; Mueller *et al.*, 1998; Scahill *et al.*, 2003) as well as progressive accumulation of white matter changes (Wolfson *et al.*, 2005).

Using VBM Good *et al.* were able to image a large cohort of healthy volunteers and compare patterns of atrophy across individuals (Good *et al.*, 2001). They observed a linear increase in the CSF compartments with age. Regionally atrophic changes were more prominent in the frontal and parietal cortices than the temporal and occipital cortices. There was also an apparently linear loss of grey matter volume asymmetrically in both parietal lobes, pre- and post-central gyri, insula, anterior cingulate cortex and left middle frontal gyrus. They described relative preservation of amygdalae, hippocampi and entorhinal cortices although it was noted that subtle changes may not have been identified due to poor alignment of the medial temporal lobe substructures.

The axiom of "a lack of statistically significant difference does not mean there is not a difference...," is very relevant for voxel-based measures assessing "t" or "p" values.

With cross-sectional imaging, age-related changes could only be inferred as different individuals were being studied at multiple time points. Accordingly, there was potential for confounding age and cohort effects. However longitudinal studies allow the use of a subject as their own control, and therefore circumvent these confounding factors. As in the cross-sectional studies, global grey matter loss has been demonstrated in elderly individuals with medial temporal lobe, prefrontal cortex, and parietal changes detected with advancing age (Tisserand *et al.*, 2004).

As well as locating regions of atrophy, these longitudinal studies provided information on how fast these changes occur. Ventricular size has been reported to increase by around 1.5cm^3 per year in individuals with a mean age of 70 and a range of 59-85 yrs (Resnick *et al.*, 2000).

Longitudinal work has shown that with increasing age there is an acceleration in atrophy rates, with the increase in rates of tissue loss being most noticeable after 70 years of age (Scahill *et al.*, 2003). Rates of global atrophy of 0.2% a year at age 30-50 and of 0.4% at age 70-80 have been reported in very healthy subjects (Fox *et al.*, 2000; Jack *et al.*, 2000; Jack *et al.*, 2004; Wang *et al.*, 2002). Likewise hippocampal atrophy rates increase from around 0.1–0.2% a year in those aged 30–50, to 0.8% in those in their mid-70s, rising further to 1.5–2% a year at 80–90 (Fox and Schott, 2004; Jack *et al.*, 2000). These more elderly cohorts may include a number of individuals with early AD (Mueller *et al.*, 1998). Both this pattern and non-linearity of atrophy has been supported by work using cortical mapping techniques (Sowell *et al.*, 2004).

In summary there are both grey matter and white matter volume losses with ageing, with regional gray matter loss more prominent in the frontal and parietal rather than other cortical regions. The relationship between ageing and cerebral atrophy is non-linear, with certain regions of the brain changing at different times to others.

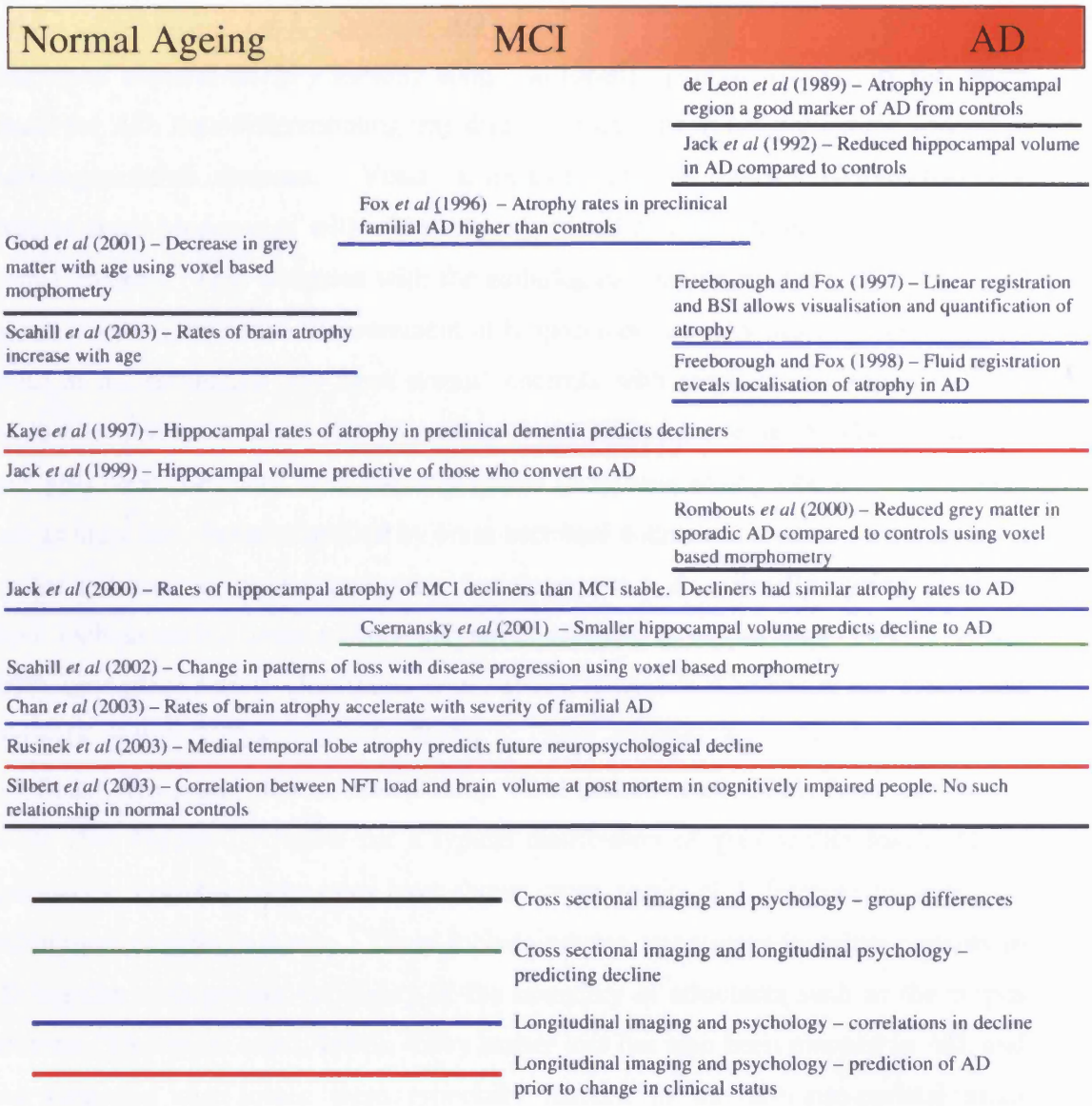


Figure 1.8 Diagram showing spectrum of disease and papers demonstrating imaging features of the disease.

1.4.2. Sporadic AD

Imaging of cerebral atrophy initially sought to identify patterns of atrophy that were typical for AD, thus differentiating this disease process from normal ageing and other neurodegenerative diseases. Visual assessment of MR images characteristically revealed small hippocampi with attendant generalised cerebral atrophy, compared with normal controls. This accorded with the pathological staging of AD progression (see below). Semi-quantitative measurement of hippocampal atrophy has been shown to be useful at discriminating AD from normal controls with some studies reporting high sensitivities and specificities (see (Scheltens *et al.*, 2002) for a review). These measures have also been correlated with memory scores (Scheltens *et al.*, 1992). Observations such as these have been quantified by cross-sectional volumetric studies which revealed smaller volumes in many structures when compared with normally ageing controls. These include medial temporal lobe sub-structures such as hippocampi (Jack, Jr. *et al.*, 1992), entorhinal cortex (Juottonen *et al.*, 1998), amygdala (Cuenod *et al.*, 1993), and cingulate gyrus (Killiany *et al.*, 2000). These changes have been mirrored by significant grey matter losses shown using VBM (Baron *et al.*, 2001; Rombouts *et al.*, 2000). (See Figure 1.9 below for a typical distribution of grey matter loss). High-dimensional mapping techniques have shown cross–sectional differences between AD and normally-ageing subjects. These include greater asymmetry in sulcal patterns in AD together with greater variability of the boundary of structures such as the corpus callosum (Thompson *et al.*, 1998). Grey matter loss has also been mapped in AD, and this suggested that losses were especially marked in the temporo-parietal areas (Thompson *et al.*, 2001).

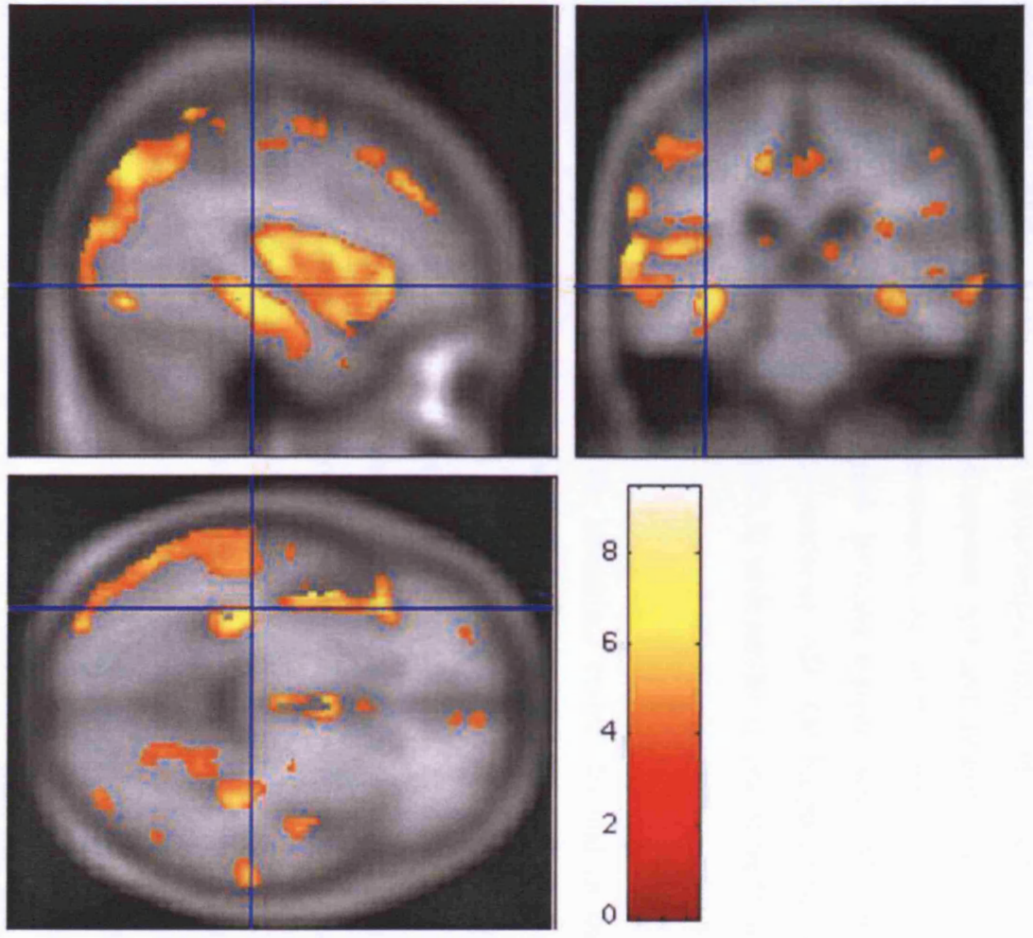


Figure 1.9 Cross-sectional VBM showing typical pattern of grey matter loss in AD compared with controls showing temporal lobe (including hippocampus) and cortical involvement in the disease process.

Medial temporal lobe substructures have been noted to be different from healthy controls at a very early stage of clinical decline. Accordingly many studies have sought to elucidate whether measures of atrophy in these regions could be used to identify AD in its earliest stages.

MRI studies in subjects with mild AD have consistently demonstrated reduced hippocampal volumes (Killiany *et al.*, 1993; Laakso *et al.*, 1995). Measurement of this area has also been shown to improve differentiation of AD from other diseases (Wahlund *et al.*, 2000). Analysis of the hippocampus using high-dimensional mapping techniques has also revealed differences between AD and control groups at one time-point and over two time-points. Cross-sectionally, the combination of surface deformity and volume has been able to distinguish between patients with AD and controls (Csernansky *et al.*, 2000). In mild to moderate AD, the hippocampal volume is typically decreased by 20-30% (see Table 1.1) with similar values being found both for amygdala and entorhinal cortex losses.

However, many have postulated that the entorhinal cortex, located in the anterior parahippocampal gyrus is the earliest structure to be affected in AD (Arnold *et al.*, 1991; Braak *et al.*, 1993). Entorhinal cortex volume losses have been found to be greater than hippocampal volume losses in mild to moderate AD compared to normal controls (Du *et al.*, 2001), although in terms of discriminating between AD and controls, entorhinal cortex and hippocampal measurements appear equally good (Juottonen *et al.*, 1999). It has been noted that patient classification significantly improves if entorhinal cortex and hippocampal measurements are combined, implying that AD is characterised by atrophy of both hippocampus and entorhinal cortex; the diagnostic advantage of greater losses in the entorhinal cortex being undermined by the variability in its measurement. Early involvement of the amygdala has been suggested by several groups (Laakso *et al.*, 1995; Lehericy *et al.*, 1994) and in some cases there was shown to be a greater difference in amydala size than parahippocampal gyrus in mild AD (Krasuski *et al.*, 1998).

Table 1.1 Cross-sectional hippocampal, entorhinal cortex and amygdala volumes in different studies.

Study Author , Year	Details (n), mean MMSE, mean age of both groups	Structure	Ratio of region (x1000) of interest volume to TIV unless indicated otherwise (SD)	% volume loss relative to controls
(Jack, Jr. *et al.*, 1992)	C (22), N/A DAT (20), N/A ≈ 75 y	Hippocampus (total)	C 2.8 (0.14‡) DAT 2.0 (0.22‡)	29‡
(Lehericy *et al.*, 1994)	C (8), 29 AD (18), 22‡ ≈ 70 y	Hippocampus (total)	C 2.38 (0.22‡) AD 1.66 (0.33‡)	30
(Laakso *et al.*, 1995)	C (16), 29 AD (32), 23 ≈ 70 y	Hippocampus	C L 3.36 (0.51) Ω C R 3.71 (0.45) Ω AD L 2.08 (0.55) Ω AD R 2.3 (0.59) Ω	33
(Krasuski *et al.*, 1998)	C (21), 30 AD (13), 24 ≈ 70 y	Hippocampus (average)	C 2.07 (0.29) AD 1.68 (0.24)	19
(Von Gunten *et al.*, 2000)	C (14) memory problems and depression (14),27+ ≈ 58 y	Hippocampus	C L 1.87 (0.31) C R 1.92 (0.26) Patients L 1.89 (0.19) Patients R 1.97 (0.17)	L 6‡ R 4‡
(Xu *et al.*, 2000)	C (30), 29 MCI (30), 26 AD (30), 21 ≈ 78 y	Hippocampus (total)	C 3.71 (0.49‡) MCI 3.25 (0.60‡) AD 2.87 (0.60‡)	MCI 12‡ AD 23‡

Study Author , Year	Details (n), mean MMSE, mean age of both groups	Structure	Ratio of region (x1000) of interest volume to TIV unless indicated otherwise (SD)	% volume loss relative to controls
(Laakso *et al.*, 2000a)	C (30), 29 AD (30), 24 ≈ 70 y	Hippocampus	C L 137.6 (23.6) * C R 149.6 (24.5) * AD L 99.5 (25.1) * AD R 101.8 (27.9) *	L 28‡ R 32‡
(Chan *et al.*, 2001b)	C (10), N/A AD (10), N/A ≈ 60 y	Hippocampus	C L1.9 (0.2) C R 2.4 (0.3) AD L 1.5 (0.2) AD R 2.0 (0.4)	L 21‡ R 17‡
(Du *et al.*, 2001)	C (40), 29 MCI (36), 26 AD (29), 18 ≈ 75 y	Hippocampus (total)	C 6.33 (0.80) TIV corrected mls MCI 5.66 (0.86) TIV corrected mls AD 4.60 (1.01) TIV corrected mls	MCI 11 AD 27
(Mega *et al.*, 2002)	C (10), 29 AAMI (10), 29 ≈ 70 y	Hippocampus	C L 2.09 (0.36) Φ C R 2.24 (0.53) Φ AAMI L 1.89 (0.30) Φ AAMI R 1.81 (0.26) Φ	L 10‡ R 19‡
(Marquis *et al.*, 2002)	C (60), 28 QD (48), 28 PCI (38), 28 ≈ 80 -86 y	Hippocampus	C 1370.3 (178.9) mm^3 QD 1242.5 (189.4) mm^3 PCI 1240.0 (190.8) mm^3	QD 9‡ PCI 10‡

Study Author , Year	Details (n), mean MMSE, mean age of both groups	Structure	Ratio of region (x1000) of interest volume to TIV unless indicated otherwise (SD)	% volume loss relative to controls
(Killiany *et al.*, 2002)	C (28), 29 questionables (73), 29 AD (16), 24 ≈ 70 y	Hippocampus (total)	C 3.54 (0.05¥) questionables 3.81 (0.06¥) AD 3.19 (0.04¥)	questionables -8‡ AD 10‡
(Lehericy *et al.*, 1994)	C (8), 29 AD (18), 22‡ ≈ 70 y	Amygdala	C 1.77 (0.25‡) AD 1.12 (0.17‡)	37
(Laakso *et al.*, 1995)	C (16), 29 AD (32), 23 ≈ 70 y	Amygdala	C L1.85 (0.23) Ω C R 1.69 (0.24) Ω AD L1.51 (0.46) Ω AD R 1.46 (0.50) Ω	14-18
(Krasuski *et al.*, 1998)	C (21), 30 AD (13), 24 ≈ 70 y	Amygdala	C 1.35 (0.24) AD 0.91 (0.24)	33
(Von Gunten *et al.*, 2000)	C (14), N/A memory problems and depression (14), 27+ ≈ 58 y	Amygdala	C L 1.31 (0.10) C R 1.37 (0.13) Patients L 1.18 (0.19) Patients R 1.27 (0.16)	L 15‡ R 13‡
(Chan *et al.*, 2001b)	C (10), N/A AD (10) , N/A ≈ 60 y	Amygdala	C L 1.21 (0.13) C R 1.32 (0.17) AD L 0.96 (0.10) AD R 1.05 (0.17)	L 21‡ R 20‡

Study Author , Year	Details (n), mean MMSE, mean age of both groups	Structure	Ratio of region (x1000) of interest volume to TIV unless indicated otherwise (SD)	% volume loss relative to controls
(Laakso *et al.*, 2000a)	C (30), 29 AD (30), 24 ≈ 70 y	Entorhinal cortex	C L 81.3 (20.0)* C R 86.2 (21.9)* AD L 58.6 (17.5)* AD R 62.9 (21.2)*	L 28‡ R 23‡
(Xu *et al.*, 2000)	C (30), 29 MCI (30), 26 AD (30), 21 ≈ 78 y	Entorhinal cortex (total)	C 0.56 (0.16‡) MCI 0.44 (0.11‡) AD 0.35 (0.11‡)	MCI 21‡ AD 38‡
(Du *et al.*, 2001)	C (40), 29 MCI (36), 26 AD (29), 18 ≈ 75 y	Entorhinal cortex (total)	C 2.73 (0.61) TIV corrected mls MCI 2.39 (0.63) TIV corrected mls AD 1.662 (0.50) TIV corrected mls	MCI 13 AD 39
(Chan *et al.*, 2001b)	C (10), N/A AD (10), N/A ≈ 60 y	Entorhinal cortex	C L 0.24 (0.05) C R 0.25 (0.07) AD L 0.17 (0.04) AD R 0.15 (0.04)	L 29‡ R 40‡
(Killiany *et al.*, 2002)	C (28), 29 questionables (73), 29 AD (16), 23 ≈ 70 y	Entorhinal cortex (total)	C 0.030 (0.01¥) questionables 0.021 (0.01¥) AD 0.012 (0.004¥)	questionables 30‡ AD 60‡

Key

L left
R right
total = left plus right
AD Alzheimer's disease
MCI mild cognitive impairment
C control
AAMI age associated memory impairment
DAT dementia of Alzheimer's type
Questionables questionable AD
QD questionable dementia
PCI persistent cognitive impairment
MMSE Mini Mental State Examination
Ω head size correction made by dividing volume by brain area
*head size correction made by dividing volume by intracranial area
Φ no head size correction
‡ calculated from information given in the paper
¥ reported as SD but more compatible with standard error of the mean
y years old
N/A not available

Although measures of entorhinal cortex atrophy provide an early and specific marker of AD, measurement reproducibility can be affected by methodological difficulties such as reliably distinguishing the entorhinal cortex region, often made more problematic by poor grey or white matter differentiation on scans. Measurements of hippocampal change are sensitive as an imaging marker of AD and with the development of semi-automated measurement techniques, often simpler to identify and measure (Haller *et al.*, 1996). Although a relatively sensitive marker of AD, hippocampal atrophy is not specific to Alzheimer's disease as significant atrophy of the hippocampus occurs in DLB, vascular dementia and FTLD (especially anteriorly in FTLD) (Barber *et al.*, 2000; Chan *et al.*, 2001b; Du *et al.*, 2002).

Volumetric longitudinal MR studies in control and AD group comparisons have allowed changes and rates of change in cerebral structures to be related to clinical stage and monitored in parallel. These studies have demonstrated an increase in whole brain atrophy rates (Fox and Freeborough, 1997) of 2-3% per year which is several times greater that that seen in age-matched controls (see Table 1.2). Rates of brain atrophy are not linear throughout the disease. In AD rates of brain atrophy have been shown to increase gradually over several years before the onset of symptoms and to accelerate as the disease progresses (Chan *et al.*, 2003). In this study the mean yearly loss of brain volume rose by an average of approximately 0.3% per year (95%CI 0.15–0.50) before subjects became symptomatic, reaching a rate of loss of over 2% per year (95%CI 2.3-3.3) by the time the mini mental state examination (MMSE) had fallen to 23/30. VBM has been used to establish patterns of tissue loss with differing disease severity, demonstrating a shift in the anatomical distribution of atrophy with disease progression in a mixed group of FAD and SAD subjects (Scahill *et al.*, 2002) (see Figure 1.10). Early AD subjects have increased rates of hippocampal (Jack *et al.*, 1998) and entorhinal cortex atrophy (Du *et al.*, 2004). Increased inward deformation of the hippocampal surface and decreased volume over time has also been shown longitudinally using high deformation techniques with rates of change similar to that found by manual volumetry (Wang *et al.*, 2003). In one study of early AD the entorhinal cortex atrophy rate was measured at 7.1+/-3.2% per year, slightly higher than that of the hippocampus (5.9+/- 2.4% per year) (Du *et al.*, 2004). Cognitive impairment correlated significantly with atrophy rates of both entrohinal cortex and hippocampus. As with the cross-sectional work, atrophy rates of the hippocampus and entorhinal cortex were comparable in differentiating between AD and controls (see Table 1.3).

Rates of loss seem to vary between studies probably reflecting both differences in measurement techniques and subject groups.

With disease progression, the inferolateral regions of the temporal lobes become more involved. Increased rates of parietal lobe atrophy are seen at all stages, whilst frontal lobe changes, and then finally occipital losses appear to occur later in the disease process. This grey matter loss has also been identified longitudinally using high-dimensional mapping, with loss starting in the temporal and limbic cortices and progressing to the frontal and occipital brain regions with relative sparing of the primary sensory-motor areas. This loss also correlates with decline in neuropsychology scores (Thompson *et al.*, 2003). Infra-tentorial changes have been studied but appear to be minor relative to cortical losses.

By comparing post-mortem findings to previously obtained MR images histopathological changes can be directly related to cerebral atrophy (Silbert *et al.*, 2003). In cognitively impaired individuals, but not in normal controls, there is a relationship between NFT accumulation and total brain volume as well as rate of brain volume change. There is also a strong correlation between amyloid plaque deposition and rate of increase in ventricular size. Rates of global brain atrophy were reported again as around 2% per year in AD with this excess atrophy mirroring the accumulation of AD pathology found at post-mortem.

There has been much interest in the predictive value of both cross-sectional and longitudinal imaging of the brain at a stage before a diagnosis of AD can be made clinically. Medial temporal lobe substructures have shown positive predictive value in identifying those who will go on to decline to AD from MCI or control status (Jack *et al.*, 1999; Jack *et al.*, 2000; Killiany *et al.*, 2000). Identification of predictive factors has been the subject of much research in both AD and MCI and is more fully addressed in the following sections.

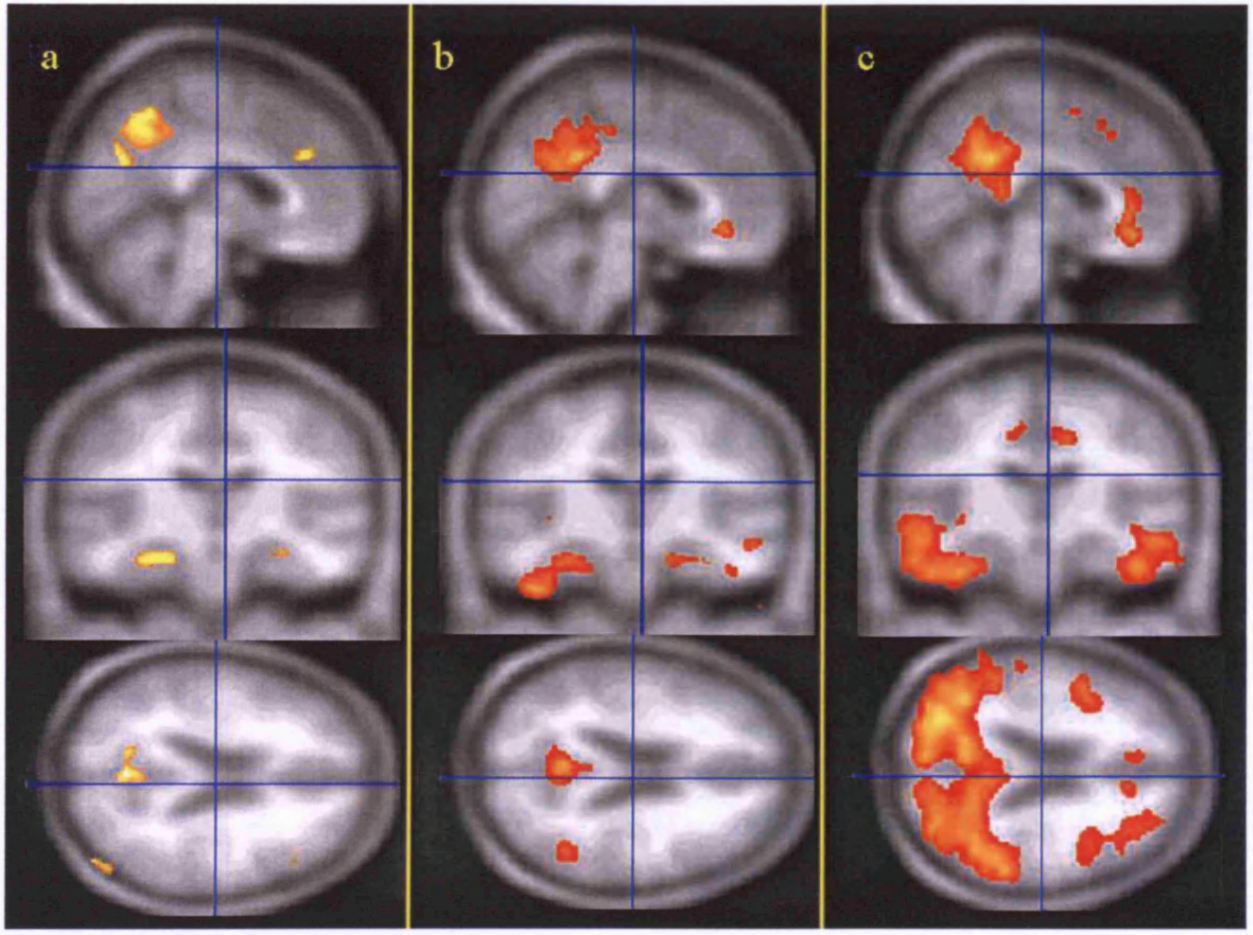


Figure 1.10 Longitudinal VBM in subjects at differing stages of AD compared with controls showing a shift in atrophy from hippocampus to neocortex as the disease progresses from a) presymtomatic b) mild and c) moderate AD.

Table 1.2 Table of whole brain rates of change in AD, MCI and normal ageing

Author, year	Subject group (n) ≈ mean age	Mean (SD) % rate of atrophy	Approximate Interval (years)
(Fox and Freeborough, 1997)	C (19) AD (9) ≈ 54 y	C 0.24 (0.32) AD 2.78 (0.92)	1
(Fox *et al.*, 2000)	C (18) AD (18) ≈ 65 y	C 0.41 (0.47) AD 2.37 (1.11)	1
(O'Brien *et al.*, 2001)	C (20) AD (9) ≈ 75 y	C 0.5 (0.7) AD 2.0 (0.9)	1
(Bradley *et al.*, 2002)	C (34) 2 with possible AD AD (5) ≈ 65-70 y	C 0.2 (0.23) AD 2.14 (0.52)	0.2-0.6
(Wang *et al.*, 2002)	C (14) AD (14) ≈ 70 y	C 0.4 (0.5) AD 2.4 (1.2)	1
(Jack *et al.*, 2004)	C stable (40) C converter (15) MCI stable (15) MCI converter (26) AD slow progressor (31) AD fast progressor (33) ≈ 76-80 y	C stable 0.4 (0.3) C converter 0.8 (0.5) MCI stable 0.4 (0.4) MCI converter 0.8 (0.5) AD slow prog. 0.6 (0.7) AD fast prog. 1.4 (1.1)	1-4
(Jack, Jr. *et al.*, 2005)	C (91) MCI (72) ≈ 80 y	C 0.5 (1.0) MCI 0.7 (0.7)	1

Author, year	Subject group (n) ≈ mean age	Mean (SD) % rate of atrophy	Approximate Interval (years)
(Schott *et al.*, 2005)	C (19) AD (38) ≈ 70 y	C 8.1 (5.0) mls/year AD 20.8 (11.2) mls/year AD 2.2 (1.2)	1
(Kaye *et al.*, 2005)	C (88) Very mild AD (23) Mild AD (27) Moderate AD (17) ≈ 75-80 y	C 0.49 (1.4) Very mild AD 1.2 (1.7) Mild AD 2.0 (2.2) Moderate AD 2.5 (2.0)	1.5-2.5

Key
AD Alzheimer's disease
C control
MCI mild cognitive impairment
y years old

Table 1.3 Table of hippocampi rates of change in AD, MCI and normal ageing

Study Author, Year	Details Subject group (n), ≈mean age y	Annualised loss as % Mean (SD)	Interval (years)
(Kaye *et al.*, 1997)	C (18) Pre-clin. dementia (12) ≈ 85-90 y	C 2.09 Pre-clin dementia 2.33	3-4
(Jack *et al.*, 1998)	C (24) AD (24) ≈ 80 y	C 1.55 (1.38) AD 3.98 (1.92)	2
(Laakso *et al.*, 2000b)	C (8) AD (27) ≈ 70 y	C 3.6 (15.1)Δ AD 7.2 (20.1)Δ	3
(Jack *et al.*, 2000)	C (58) ≈ 80 y MCI (43) ≈ 77 y AD (28) ≈ 74 y	C 1.9 (1.1) MCI 3.0 (1.6) AD 3.5 (1.8)	3
(Cardenas *et al.*, 2003)	C (16) ≈ 76 y CI (6) ≈ 70 y AD (7) ≈ 76 y	C 1.8 (0.8) CI 1.9 (2.3) AD 5.4 (2.8)	2-3
(Jack *et al.*, 2004)	C stable (40) C converter (15) MCI stable (15) MCI converter (26) AD slow progressor (31) AD fast progressor (33) ≈ 76-80 y	C stable 1.4 (1.2)Ψ C converter 3.3 (2.4)Ψ MCI stable 1.8 (1.7)Ψ MCI converter 3.3 (2.9)Ψ AD slow prog. 3.0 (4.5)Ψ AD fast prog. 3.6 (3.2)Ψ	1-4
(Du *et al.*, 2004)	C (25) AD (20) ≈ 75 y	C 0.8 (1.7) AD 5.9 (2.4)	2

Study Author, Year	Details Subject group (n), ≈mean age y	Annualised loss as % Mean (SD)	Interval (years)
(Hashimoto *et al.*, 2005)	AD (93) ≈ 70 y	AD 5.04 (2.54)	1
(Jack, Jr. *et al.*, 2005)	C (91) MCI (72) ≈ 80 y	C 1.7 (1.4) MCI 3.3 (2.7)	1
(Thompson *et al.*, 2004a)	C (14) AD (17) ≈ 70 y	C 2.1 (2.8) AD 6.8 (10.2)*	1-3
(Jack *et al.*, 2003)	AD (192) ≈ 73 y	AD 4.9 (-0.5 – 15.2)#	1
(Wang *et al.*, 2003)	C (26) DAT (18) ≈ 74 y	C L 4.0 R 5.5Δ DAT L 8.3 R 10.2Δ	2
(Fox *et al.*, 2005)	AD placebo (57) ≈ 70 y	AD placebo 2.86 (3.19)Δ	> 1
(Kaye *et al.*, 2005)	C (88) Very mild AD (23) Mild AD (27) Moderate AD (17) ≈ 80 y	C 2.2 (6.0) Very mild AD 4.3 (7.4) Mild AD 2.9 (7.8) Moderate AD 3.2 (6.8)	2
(Mori *et al.*, 2002)	AD APOE ε4+ (38) AD APOE ε4- (17) ≈ 70 y	AD ε4+ 9.76(4.27) AD ε4- 6.99 (4.24)	1

Key
y approximate mean age of subjects in study in years. Given for different subject groups if these appear different.
AD Alzheimer's disease
C control
MCI mild cognitive impairment
DAT dementia of Alzheimer Type
CI cognitive impairment
Pre. clin: pre-clinical
† results presented here are 1/3 of those quoted in paper as change was presented as loss over a three year period.

*given by personal communication
median and range
Δ percentage loss over interval (i.e. not annualised)
APOE Apolipoprotein gene

Table 1.4 Table of ventricular rates of change in AD, MCI and normal ageing

Study Author, Year	Details Subject group (n), ≈ mean age y	Annualised loss Mean (SD)	Interval (years)
(Jack *et al.*, 1998)	C (24) AD (24) ≈ 80 y	C 6.15% (7.69) ^ AD 14.16% (8.47) ^	2
(Wang *et al.*, 2002)	C (14) AD (14) ≈ 70 y	C 0.79 ml AD 8.20 ml	1
(Wang *et al.*, 2002)	C (14) AD (14) ≈ 70 y	C 1.9% (4.2) AD 13.8% (4.8)	1
(Bradley *et al.*, 2002)	C (32) 2 with possible AD AD (5) ≈ 65-70 y	C 4.1% (0.9) AD 13.0% (2.4)	0.2-0.6
(Silbert *et al.*, 2003)	C (8) AD (20) over 80 y	C 3.3 ml (3.5) AD 5.5 ml (3.2)	4
(Jack, Jr. *et al.*, 2005)	C (91) MCI (72) ≈ 80 y	C 2.4% (2.0) MCI 3.3% (2.3)	1
(Kaye *et al.*, 2005)	C (88) Very mild AD (23) Mild AD (27) Moderate AD (17) ≈ 80 y	C 3.5% (3.8) V. mild AD 7.4% (4.4) Mild AD 9.9% (5.8) Mod. AD 12.1% (4.8)	2

Key
y approximate mean age of subjects in study in years. Given for different subject groups if these appear different.
AD Alzheimer's disease
C control
MCI mild cognitive impairment
^ temporal horns only

Table 1.5 Table of entorhinal cortex rates of change in AD, MCI and normal ageing

Study Author, Year	Details Subject group (n) ≈mean age y	Annualised loss Mean (SD)	Interval (years)
(Cardenas *et al.*, 2003)	C (16) ≈ 76 y CI (6) ≈ 70 y AD (7) ≈ 76 y	C 2.6% (2.5) CI 6.5% (4.8) AD 9.3% (4.4)	2-3
(Jack *et al.*, 2004)	C stable (40) C converter (15) MCI stable (15) MCI converter (26) AD slow progressor (32) AD fast progressor (33) ≈ 76-80 y	C stable 2.9 (2.6)Ψ C converter 5.1 (5.1)Ψ MCI stable 3.7 (3.7)Ψ MCI converter 6.8 (4.4)Ψ AD slow prog. 8.0 (5.8)Ψ AD fast prog. 8.4 (9.2)Ψ	1-4
(Du *et al.*, 2004)	C (25) AD (20) ≈ 75 y	C 1.4 (2.0) AD 7.1 (3.2)	2
(Jack, Jr. *et al.*, 2005)	C (91) MCI (72) ≈ 80 y	C 5.0 (3.6) MCI 7.0 (4.3)	1

Key
y approximate mean age of subjects in study in years. Given for different subject groups if these appear different.
AD Alzheimer's disease
C control
MCI mild cognitive impairment
CI cognitive impairment
Ψ medians (inter-quartile ranges)

1.4.3. Familial AD

MRI studies carried out in populations at risk of developing Alzheimer's disease owing to inheritance of presenilin PS1 or PS2 or APP mutations have led to characterisation of familial AD and to the defining of a 'preclinical' phase of dementia. Serial MR scanning of individuals within 5-8 years of the historical onset of the disease showed that significant hippocampal atrophy was detectable on MRI at an early stage in the disease when subjects were still apparently asymptomatic (Fox *et al.*, 1996b). Registration of serial MRI in these FAD subjects demonstrated that rates of atrophy increased at an early stage prior to the onset of measurable cognitive deficit, with a median rate of global atrophy of 1.0% per year (95% CI 0.78–1.88) (Fox et al., 2001). Schott *et al.* measured medial temporal lobe atrophy in their series at 5.7 % per year (95% CI 3.62 -7.83). Extrapolation of these rates of atrophy back in time suggested that medial temporal lobe atrophy might have commenced at least 3.5 years prior to onset of symptoms (Fox *et al.*, 2001; Schott *et al.*, 2003). These findings clearly outline the disease continuum that is present in Alzheimer's disease from a presymptomatic phase to cognitive decline sufficient to fulfil criteria for AD. Individuals with familial Alzheimer's disease have an earlier age of onset and often different clinical features to those with sporadic AD and therefore care must be taken in relating these results to sporadic AD directly (Rossor *et al.*, 1996). However a similar picture has been described using serial MR imaging in the early stages of sporadic AD as described below.

1.4.4. Amnestic MCI and presymptomatic sporadic AD

Comparison of studies performed in this group of patients is complicated by the wide number of diagnostic criteria used for MCI subjects in imaging studies. These have ranged from isolated memory impairment to one domain (non-memory) cognitive impairment or multi-domain mild cognitive impairment. The proportion of AD subjects may therefore vary widely between studies with 'MCI' groups comprising of variable amounts of 'worried well' subjects and non-AD pathology.

Research groups have increasingly recognised the MCI subjects as important in terms of diagnosis and treatment and therefore have sought to identify the earliest imaging markers in MCI much like the approach in AD. In 1989 de Leon *et al.* published an important study showing that qualitative estimates of hippocampal atrophy in memory-impaired subjects predicted decline to AD (de Leon *et al.*, 1989). Later studies also

have shown that in subjects with MCI, small baseline hippocampal size is predictive of subsequent conversion to AD (Jack *et al.*, 1999).

Du *et al.* found that entorhinal cortex and hippocampal volume measurements were significantly reduced in MCI compared to controls (Du *et al.*, 2001). The magnitude of entorhinal cortex atrophy was similar to that of hippocampal atrophy in MCI. Volume losses were not as great as were seen in AD. Killiany *et al.*, in a three year follow-up study of patients with varying degrees of cognitive impairment, found that overall classification into MCI and control groups with the entorhinal cortex was better than 85% (Killiany *et al.*, 2002). However, Du *et al.* found that the hippocampus was better than the entorhinal cortex in distinguishing MCI from controls and even adding the entorhinal cortex to the hippocampus did not improve classification (Du *et al.*, 2001). Xu *et al.* found that the entorhinal cortex and hippocampus had equivalent discrimination power between MCI and controls and between AD and MCI (Xu *et al.*, 2000). This suggested that, as in AD, the entorhinal cortex offers little practical advantage over the hippocampus in differentiating MCI from controls. The relative merits of hippocampal and entorhinal cortex measures remains controversial with advocates for both structures.

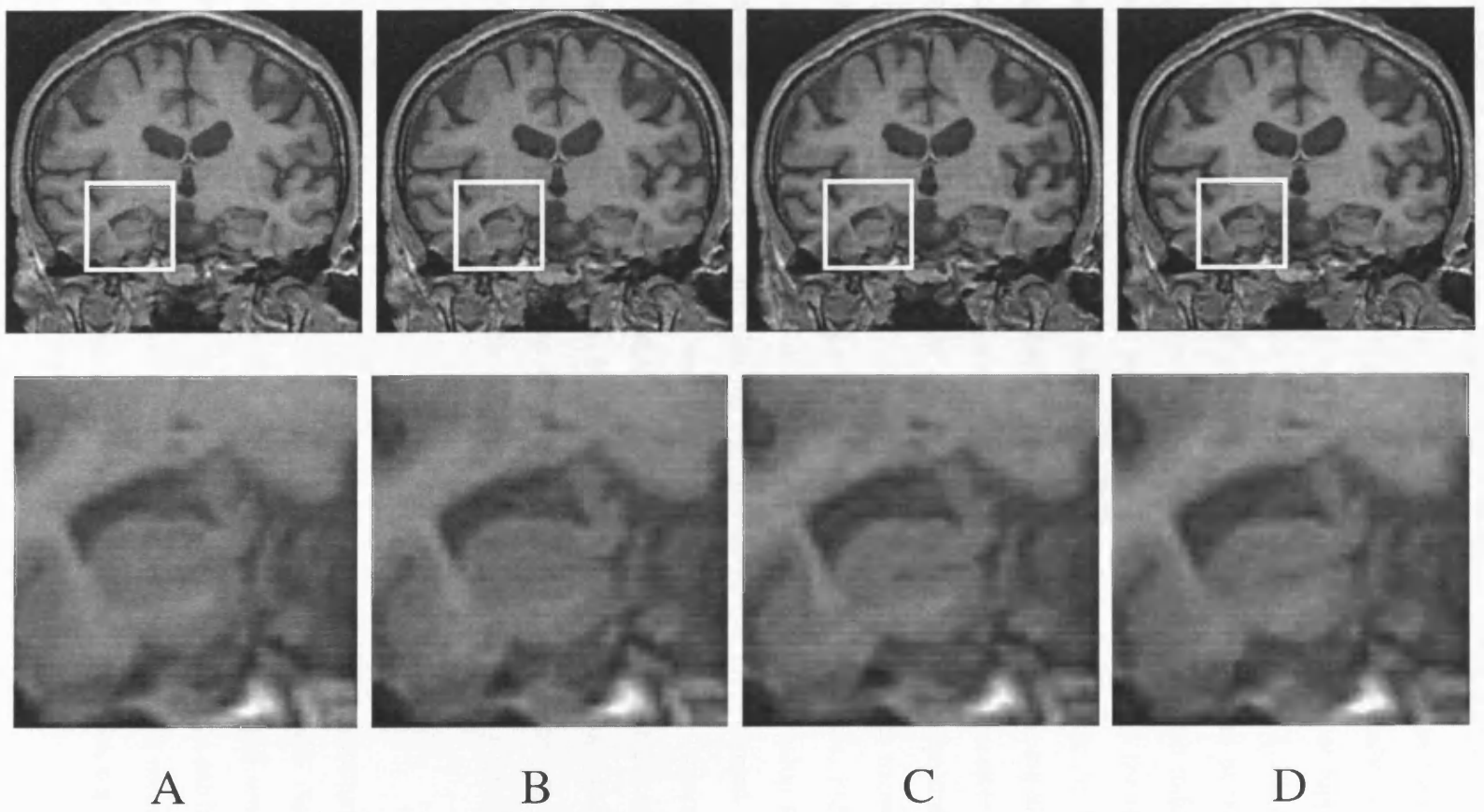


Figure 1.11 Demonstration of hippocampal shrinkage over time in an individual who at presentation and nine month interval had subjective memory complaints.
Time-points A: baseline, B: nine months after baseline imaging, C: two years after baseline imaging, and D: three years after baseline imaging.
By the final scan D the individual had progressed to fulfill criteria for AD. The individual had subsequent post-mortem confirmation of AD.

Longitudinal studies have shown that MRI measures of medial temporal lobe atrophy combined with neuropsychometric tests have a high predictive value for progression to AD (Arnaiz and Almkvist, 2003; Petersen *et al.*, 2001b). One study (Jack *et al.*, 2000) demonstrated that over a period of three years significantly greater hippocampal atrophy had occurred in a group of MCI individuals who progressed to a diagnosis of AD than those who remained cognitively stable. Interestingly a proportion of their control group also deteriorated cognitively over the follow-up period to become rediagnosed as MCI. They had a significantly higher rate of hippocampal atrophy than the rest of the control group. This suggests that a prodromal phase also exists prior to MCI in sporadic Alzheimer's disease where cerebral atrophy and specifically hippocampal atrophy is taking place but definite cognitive impairment may not be clinically detectable (see Figure 1.11 for an example). Rates of control-converter subjects (who started the study as controls but at follow-up had AD) and MCI-converter subjects were similar to those of AD, whereas stable MCI rates were similar to those of stable normal subjects, indicating that MRI measures of change correlated more closely with disease progression than with clinical assessment at baseline. This also implies that the rate of hippocampal atrophy increases prior to the MCI stage but then seems to stay relatively constant (at around 3-4% per year) as MCI progresses to established AD. Similarly, another longitudinal study (Rusinek *et al.*, 2003), demonstrated that medial temporal lobe (MTL) atrophy was able to predict which individuals with normal baseline cognitive functioning would convert to a diagnosis of MCI. The overall accuracy of the MTL prediction was 89% with a specificity of 94% and sensitivity of 77%.

There has been some debate as to whether medial temporal lobe structures or measures of whole brain atrophy or ventricular size most closely correlate with disease progression. One group compared hippocampi, entorhinal cortex, ventricles and whole brain atrophy rates to address this question. They found whole brain atrophy and ventricle size correlated most strongly with disease progression. They also found their MR measurements aligned better with disease progression (subjects changing clinical diagnoses, e.g. MCI to AD) than their measures of neuropsychometric change (Jack *et al.*, 2004).

Volumetric measures are free from subjective effects such as motivation, general health, anxiety, depression or fatigue that may contribute to cognitive test and rating scale variability. Where it is uncertain if an individual reaches criteria for mild cognitive impairment or other diagnosis, MRI provides an objective measure of disease, adding clinical certainty and perhaps giving information regarding prognosis.

1.5. Conclusion

With the potential development of disease modifying therapies, parallel development and assessment of surrogate markers of disease progression for use in clinical trials is increasingly important. Studies have suggested that change in MRI atrophy rates may correlate better with disease progression than change on cognitive tests or clinical rating scales (Jack *et al.*, 2004). Measures of atrophy rates for different areas of the brain using different techniques are also being evaluated so that the most suitable ones may be used in such trials. One such study which aims to assess this is the Alzheimer's disease neurobiology initiative (ADNI): website http://www.loni.ucla.edu/ADNI. A major aim of such studies is to assess whether the addition of different MRI measures of change to therapeutic trials will allow smaller sample numbers or shorter, more efficient trials.

Major progress has been made in quantitative imaging of the brain over the last ten years. Advances in imaging and research tools have proved useful clinically and have allowed better characterisation of the disease process. Currently, structural imaging is recommended as part of the diagnostic process to differentiate other pathology from neurodegenerative causes of cognitive impairment. It also provides a useful adjunct to both neuropsychometric and clinical assessments in the differential diagnosis and in monitoring disease progression. In the future, structural imaging may provide useful prognostic information as well as acting as a potential surrogate marker for disease progression or regression in clinical trials for AD. There is a requirement for greater automation of structural imaging techniques in order to evaluate the number of scans taken clinically in addition to those from clinical trials.

2. REGISTRATION, SEGMENTATION AND USE OF TEMPLATES

2.1. Chapter introduction

Scan registration and segmentation are essential for many imaging analyses. They have enabled research in neuroimaging to progress from visual assessments to automated measurement techniques where images from hundreds of patients can be assessed simultaneously. Registration refers to the aligning of one or a number of images so that they are in the same spatial framework. Segmentation is the delineation of structures or tissue types within an image. Registration and segmentation are intrinsically linked; many forms of registration require some form of segmentation to enable the registration procedure to be successful and to ensure the images are aligned over regions of interest. Segmentation of some structures is best performed following registration to ensure that similar arbitrary decisions are made for all subjects. Registration and segmentation can combine to make powerful tools to assess MR images. They are able to localise and quantify differences between subjects or within a subject over time. Application of these techniques has moved imaging in dementia from the exclusion of other possible causes to become a predictive tool with diagnostic utility (Scheltens *et al.*, 2002).

Segmentation of structures within the brain is important both for diagnosis and as a marker of disease progression. Whole brain atrophy may be common to many diseases but different diseases have distinctive patterns of atrophy in specific brain regions. As a result, segmentation of different areas of the brain has been employed in order to i) understand the natural history of the disease and ii) develop reliable non-invasive markers of the disease. Many of the "gold standard" methods of segmentation are time consuming and therefore techniques which are able to increase the automation of such measurements improve the likelihood of such a technique being widely applicable. In this chapter, the basis of developing reliable markers and increased automation of these markers using registration and segmentation will be discussed.

2.2. *Intra-subject registration of images*

To improve the visual assessment of serial images, scans can be co-registered using a variety of techniques with differing levels of matching. The registration process removes the effect of differing head positions of the subject in the scanner and brain positions within the skull by transforming and resampling one scan to match the other (see Figure 2.1). In addition to allowing both obvious and subtle changes to be visualised, registration allows such changes to be localised and measured. Such procedures (for example, those described in detail in (Friston *et al.*, 1995a; Jenkinson and Smith, 2001; Ostuni *et al.*, 1997; Thevenaz *et al.*, 1998; Woods *et al.*, 1998; Zhilkin and Alexander, 2000)) can be run through a number of different packages some of which are freely available.

2.3. *Linear and non-linear registration and the resulting correspondence*

Linear registration describes techniques where the parameters required to map the source scan to the target scan are applied equally to every voxel within the image. The result of this process is a resampling of the source image to the target image producing a pair of aligned images. Non-linear registration describes when registration parameters are allowed to vary throughout the image which allow for the source and target image to be more accurately matched (see Figure 2.1). The choice of which registration technique should be employed depends on the question being addressed and the validity and accuracy of the answer will depend on the registration procedure employed.

The result of any registration is that the two co-registered images are in correspondence, i.e. one point on one image corresponds to a particular point on the other image. These two points will, in theory, represent the same type of tissue in each scan. However, no measurement, or in this case scan, is completely accurate and in some applications of registration, for example image-guided neurosurgery, it is important to know what the margin of error is in order to know how to interpret the co-registered scans.

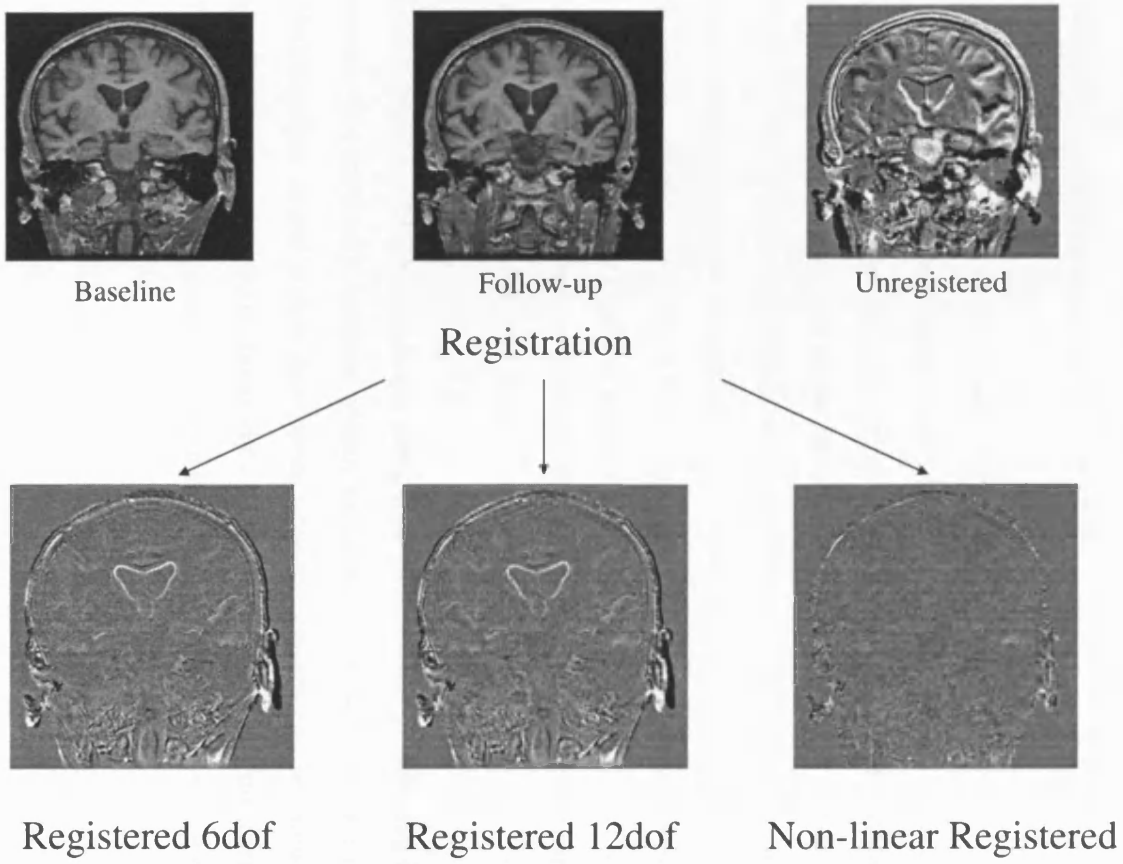


Figure 2.1 Figure to show the effect of registration methods.
Difference images (baseline image – repeat image) of the unregistered, and six degrees of freedom, affine, and non-linear registered images are shown. Registration of at least six degrees of freedom will remove positional changes. With increasing degrees of freedom (zero for unregistered to infinite for non-linear) increasing levels of matching between the baseline and registered images are achieved.

The use of registration in serial imaging of dementia, where tissue is lost between the two imaging time-points, complicates the issues of correspondence. No longer will every point between the images correspond following rigid-body registration as the atrophy that has occurred means that tissue is lost and replaced by fluid. The assumption that many atrophy measures make is that images should not change over time, and therefore the lack of correspondence between scans (often seen at the borders of the brain) is a measure of the amount of atrophy the brain has undergone.

2.4. Components of the registration procedure

Registration is achieved when transformation parameters are determined which, when applied, will match one image to another. To achieve this, registration algorithms have been developed to estimate these transformation parameters. Often the first operation which is required for registration to be performed is the determination of a region over which the images are to be matched. This may be the segmentation of a brain region or may be a method to generate a region which approximates to the inside of the skull. Next, the transformation parameters necessary to match the source to the target scan are calculated by an optimisation process. Finally, these transformation parameters are applied to the source image by interpolation.

Registration algorithms are usually composed of a number of differing components. These include: i) a similarity measure which calculates how well the two images match, ii) a transformation model which determines how the images can be transformed, and iii) an optimisation process which determines whether the similarity measure can be improved. Most registration algorithms are solved iteratively until exit criteria are met.

2.4.1. Similarity measures used for registration

The similarity of an image pair is usually based on a point, contour or voxel intensity correspondence between the images. The voxel intensity similarity can be cross-correlation, sums of squared differences, ratio image uniformity, mutual information or normalised mutual information. For a review of these differing techniques see Crum *et al.* (Crum *et al.*, 2004a) or Hill *et al.* (Hill and Batchelor, 2001).

2.4.2. *Interpolation schemes used for registration*

Interpolation is necessary to estimate transformation parameters and for image transformation as the estimates of voxel intensity values are needed for voxels between those in the original image. Once the optimum transformation parameters are obtained the source image can be resampled to match the space of the target image. A number of interpolation schemes are available for doing this. Techniques include those where values of the nearest voxel are taken, to those which take a weighted average from a number of voxels. Tri-linear interpolation generates new values based upon the mean of adjacent voxels which are weighted according to their distance from the point. This form of interpolation is prone to errors and as a result can reduce the sensitivity to detect subtle changes between scans (Hawkes, 2001). Windowed-sinc interpolation is most commonly used and has been shown to be relatively fast and accurate (Thacker *et al.*, 1999).

2.5. *Application of linear registration*

For a review of technical details of these registration techniques see Ashburner *et al.* (Ashburner *et al.*, 2003). Several automated techniques have been designed to accurately spatially match scans (Collins *et al.*, 1994; Freeborough *et al.*, 1996b; Jenkinson and Smith, 2001; Woods *et al.*, 1998). These are based upon matching voxels according to specific similarity measures such that all structures approximate to corresponding regions on each scan of a registered pair. Such registrations can be performed on the whole image or can be based upon matching the two brain areas of the scans following semi-automated brain segmentation (Freeborough *et al.*, 1997) or automated brain extraction (Smith *et al.*, 2001). Brain-brain registration means that only the two brains alone are spatially matched rather than the skulls or non-brain material. This prevents the potential changes in relative brain to skull position or chemical shift within the skull and scalp affecting the brain matching. The simplest form of registration is rigid registration which involves rotations and translations (see Figure 2.2) in the x, y and z axes of the image. This type of registration aligns scans but does not remove fluctuations in voxel size (voxel drift) due to changes in the scanner over time as it does not scale the source image to match the target. A nine degrees of freedom (dof) registration or rigid plus scaling registration is commonly used. The added scalings, in the x, y and z axes are able to remove changes in voxel size over time and this type of registration has been validated extensively (Freeborough *et al.*, 1996b). The spatial matching of the images may be improved further by adding shears in

the x, y and z planes to make a 12 dof (affine) registration. However, by increasing the level of matching, some atrophy may be removed by the registration process, leading to underestimation of atrophy.

2.6. Quantifying change in the brain following linear registration

The amount of loss that has occurred in the brain over time can be calculated by applying semi-automated procedures such as the brain boundary shift integral (BBSI) to a pair of scans. This has been validated (Freeborough and Fox, 1997) as a robust and accurate measure of brain loss. The technique is summarised below in Figure 2.3. Briefly, the BBSI calculates change at the border of the brain and quantifies this change in intensity between the two scans at this border as brain loss. This loss can be visualised by colour overlays and can give a value for loss of brain tissue. Other techniques have also been developed that measure whole brain loss in a similar way (SIENA) (Smith *et al.*, 2001). Although these techniques are able to quantify loss they are unable to determine precisely where this loss occurred. Loss visualised at the border of the brain may mean loss of grey matter directly below the grey matter-CSF border or may be a secondary effect of white matter loss at some distance from the border; usually changes at the brain surface reflect multiple contributions and readjustments to losses at a cellular level.

These methods have been used successfully to quantify rates of whole brain atrophy in normal ageing and AD, originally in (Fox and Freeborough, 1997) and more recently in comparisons carried out to evaluate differences between atrophy rate measures and clinical disease progression in AD using a modified version of SIENA (Jack *et al.*, 2004). There is much interest in these rates of atrophy as possible surrogate markers for therapeutic effect in drug trials. However there are limitations to the use of whole brain atrophy as a surrogate marker, and it may be that rates of localised atrophy provide a better indicator of disease progression. This will be discussed in more detail in the final section.

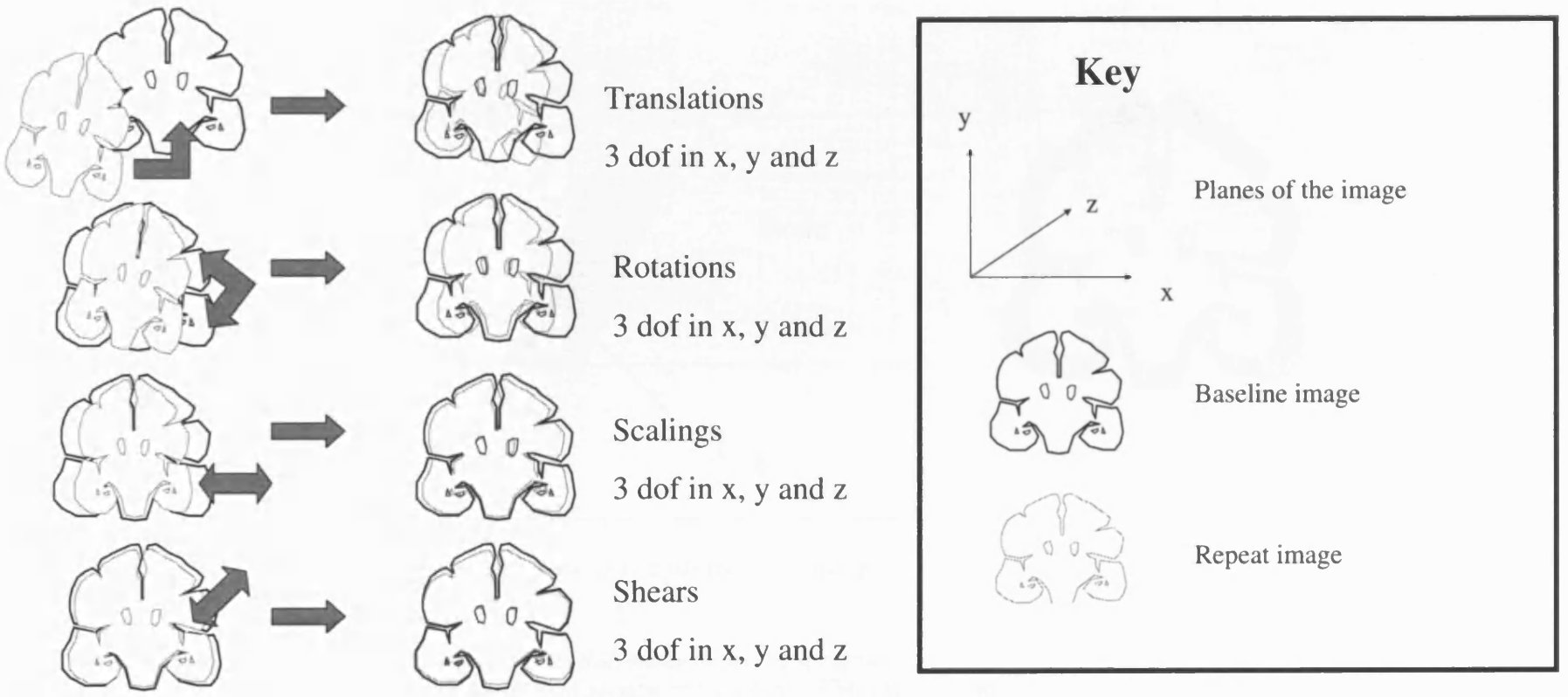


Figure 2.2 Schematic representation of registration with increasing degrees of freedom.
Numbers of degrees of freedom increase incrementally by three as each level of matching is applied in the three dimensions of the image (x, y and z).

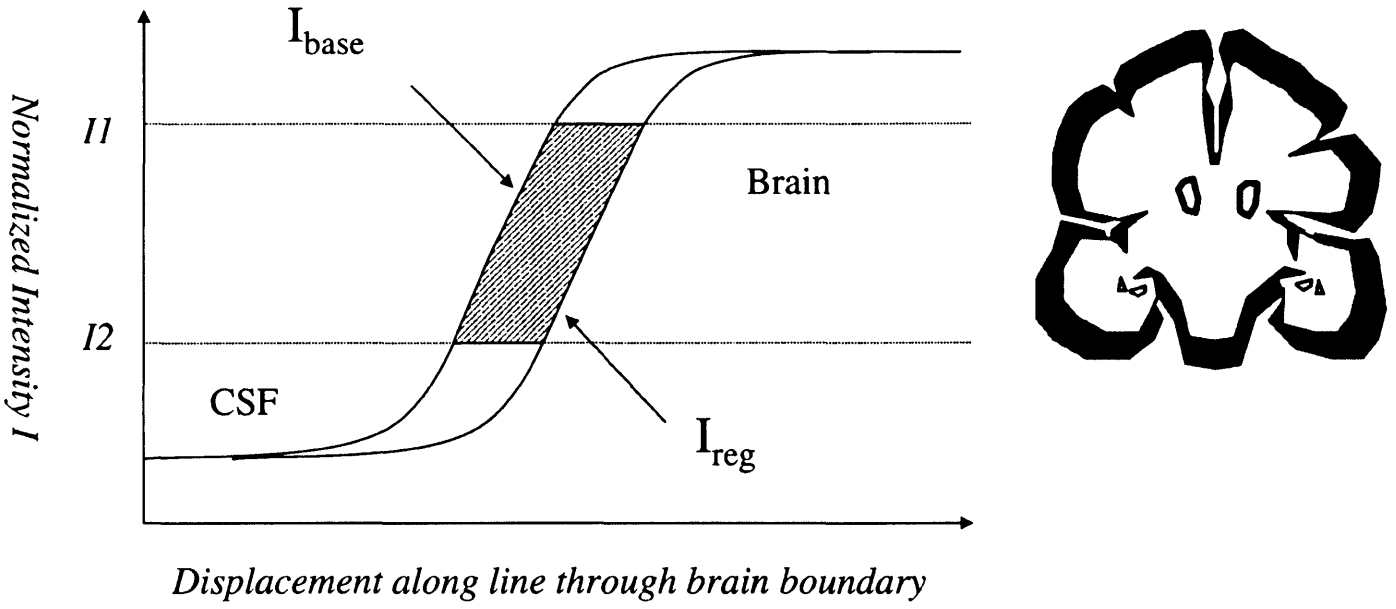


Figure 2.3 Representation of the boundary shift integral (BSI) measure in one dimension.
The boundary shift measure is represented by the hatched section of the graph. This change is measured over an area of the brain which is represented by the grey area of the brain to the right. This area is the union minus the intersection of these two registered brain regions, resulting in an area which encompasses the border of the brain.

2.7. *Non-linear registration*

Linear registration is unable to localise atrophy to specific tissues and as a result non-linear techniques have been developed not only to assess how much change occurs between scans, but from which tissues the loss occurred. Such non-linear techniques can be applied following some degree of linear spatial matching described above to remove positional differences within the brain. These techniques aim to remove all differences between the images by iterative solving of equations which model how the brain tissue intensities may have changed. The brain can be modeled, for example, as a viscous compressible fluid (Christensen *et al.*, 1996; Freeborough and Fox, 1998), or as elastic material (He and Christensen, 2003). Other forms of non-linear registration exist, for example, free form deformation based on B splines (Rueckert *et al.*, 1999; Studholme *et al.*, 2001). The result of a non-linear registration is a registered repeat image (which should match the baseline image exactly), and a deformation field which describes how each voxel should be transformed to match the source to the target. This deformation field can be used to calculate a stretch (Jacobian) file which attributes values to voxels corresponding to the amount of contraction or expansion each voxel has had to undergo for the images to match. The stretch file can be visualised as a colour overlay on the baseline image which can give information regarding where atrophy has occurred (see Figure 2.4). The values in the stretch file can also be integrated over brain areas (for example, whole brain region) which can give a value corresponding to amount of loss within that region over time.

2.8. *Bias field inhomogeneity*

Both linear and non linear registration techniques and their corresponding methods of atrophy quantification can greatly aid assessment of a disease progression. However, it is beneficial for scan pairs (target and source images) to be of similar quality for the quantifications and localisations of change to be accurate. Similar scan pair quality means similar grey matter / white matter contrast, and intensity inhomogeneity (bias) between scans. Bias is caused by a number of factors: inhomogeneity of the magnetic field of the MR system, inhomogeneity of the radiofrequency pulse or nonuniformity in the sensitivity of the receiver coils used to detect the MR signal. A number of different techniques have been used to remove this bias field on individual images, for example the intensity histogram sharpening technique: N3 (Sled *et al.*, 1998). Other techniques have been developed to differentially adjust the bias field of a scan pair such that the bias field of the

difference image is corrected to make the bias of the scan pair similar (see Figure 2.5) (Lewis and Fox, 2004). This form of bias correction will only correct for differential bias but will not correct for bias common to both images. Differential bias correction can be applied with or without individual scan bias correction.

2.9. Segmentation

Many areas of the brain are of interest to researchers investigating potential surrogate markers of disease. Analysis of regions is either cross-sectional (comparing volumes between groups) or longitudinal (comparing rates of change in structures over time). Serial volumetric analyses are best completed on either nine dof or affine registered images as these types of registration include scalings which corrects for the inevitable changes in MRI magnetic field due to scanner gradient change (Whitwell *et al.*, 2004b) or positional changes of a subject relative to fixed scanner geometry (Gunter *et al.*, 2003). Such scanner-related changes can result in fluctuation of voxel sizes which can complicate comparisons of such measures (Whitwell *et al.*, 2001). The "gold standard" of image segmentation is manual measures, however these are operator-intensive and are subject to manual errors. As a result, more automated techniques are being developed to improve the applicability of regional segmentation in a clinical scenario.

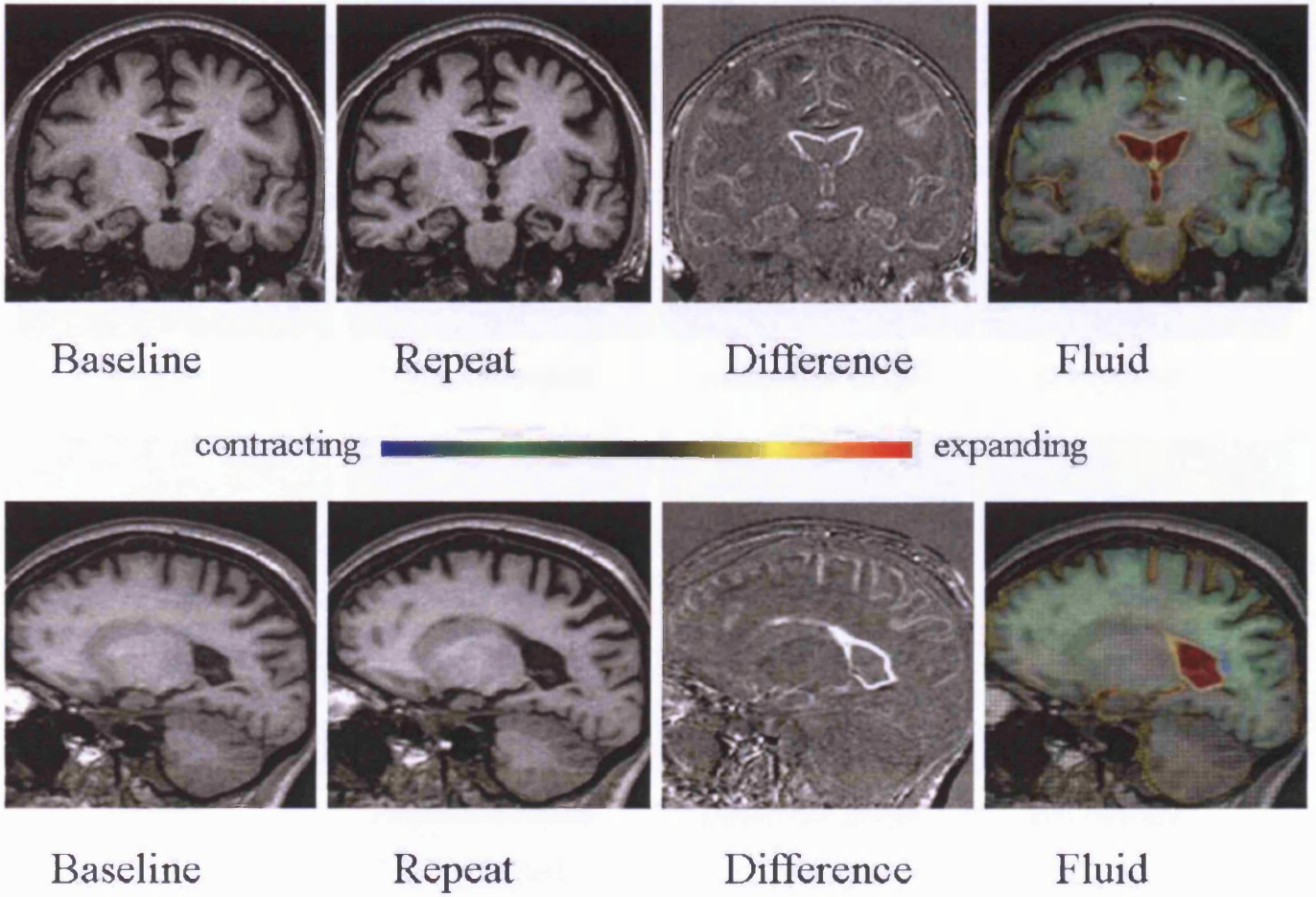


Figure 2.4 Baseline, nine dof registered repeat, difference image and stretch file created from fluid registration of the two nine dof registered scans in both control and AD over the same interval.

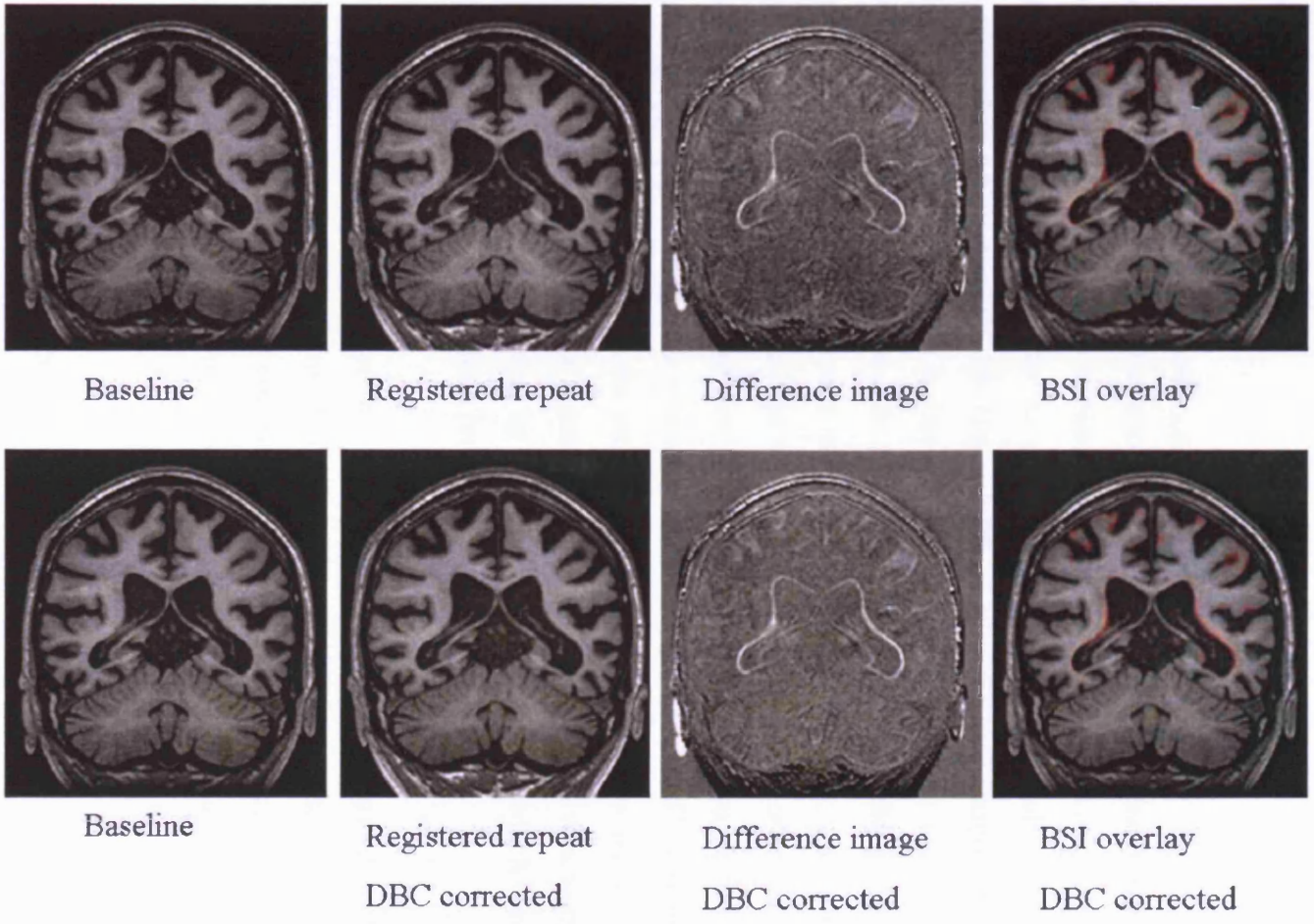


Figure 2.5 Differential bias correction (DBC) on registered images.
Top showing registration without DBC and below showing with DBC. The boundary shift integral overlay on DBC corrected images shows less artefactual cererbellar gain (green) whilst maintaining real loss (red).

2.10. Common terms used to describe aspects of segmentation

When new methods or segmentation protocols are discussed or new segmentation-based markers are developed, certain terms are used to describe characteristics of the method of segmentation.

2.10.1. Measurement of segmentation accuracy

Accuracy of segmentation is difficult to measure. This is because what is segmented on a scan is not real tissue which can be histologically characterised but is only a representation of the macroscopic brain structure. As a result there exists no true "gold standard" as there are no studies which quantify the real volume of structure and compare this with MR volume *in vivo*. One of the main problems with potentially comparing a structure on a scan with that structure at post-mortem is the fact that the brain will change in size due to changes following death. One study has imaged the brain following death and compared the volumes of the real hippocampi with volumes on the post-mortem imaged brains (Gosche *et al.*, 2001). Owing to few of these types of studies being performed and the impracticality of measuring real volume *in vivo,* the "gold standard" of segmentation is regarded to be manual segmentation and this is against what most new segmentation-based techniques are measured. Manual segmentation is also prone to errors which originate from differing interpretation of a protocol by different raters, any artefacts which may be present on the scans and the scan acquisition protocol.

Volumes of structures measured by new protocols can be compared with manually segmented volumes using a form of voxel overlap. A simple measure of intra-observer repeatability can be obtained by comparing the volumes of a structure segmented on pairs of scans acquired over a short space of time, using the following formula:

$$\text{Fractional volume difference} = \frac{2*|V1-V2|}{(V1+V2)}$$

Equation 2.1

However, random segmentation errors can result in a similar reported segmented volume obtained by labelling a different set of voxels. A more stringent measure of repeatability is the volumes of voxels which are labelled consistently in the two segmentations.

This can be determined using the following formula:

$$D = \frac{VoD}{VoD + VoS} \qquad S = \frac{VoS}{VoD + VoS}$$

Equation 2.2

In words, the voxel difference D is the fraction of inconsistently labelled voxels and the voxel similarity S is the fraction of consistently labelled voxels. VoS (the Volume of Similarity) is the volume of the intersection of the two regions; VoD (the Volume of Difference) is the volume of the union minus the intersection. For two segmentations in complete agreement D = 0 and S = 1. For two segmentations in complete disagreement (i.e. non-overlapping but which could nonetheless have the same volume) D = 1 and S = 0. For all other region pairs $0 < D < 1$. These measures are similar to those discussed by Andreasen *et al.* (Andreasen *et al.*, 1996) and have been used in one form or another by a variety of studies (see Table 2.1).

2.10.2. Measurement of reliability of segmentation

Reliability is defined as the degree to which multiple assessments of a subject agree (sometimes this is referred to as reproducibility). This can be calculated either within or between raters. Reliability can be measured a number of ways but the figure that is often quoted in the text of volumetric papers is the quantification described in Equation 2.1. However, a more standard statistical approach is the reliability co-efficient or intraclass correlation coefficient (ICC) (see below) (Bartko, 1966; Ebel, 1951; Fleis, 1986). This describes the proportion of variation that is not due to measurement error where the value has extremes of 0 and 1 where unity shows a method to be completely reliable.

$$\text{reliability coefficient} = \frac{\sigma_b^{\,2}}{(\sigma_b^{\,2} + \sigma_w^{\,2})}$$

Equation 2.3

where σ_b denotes between subject variance and σ_w denotes within subject variance

2.11. Subdivision of segmentations

As discussed earlier in section 2.3.2 there may be areas of a structure that are more variable in volume and shape compared with others or more affected by a disease process than others. To understand the natural history of a particular disease and / or improve precision of a marker, segmented structures may be subsequently subdivided. For example, the head of the hippocampus may be more reduced in volume relative to the tail in AD, therefore measuring the tail in addition to the head of hippocampus may only add noise and variability to the measure. Specific analysis of the hippocampal head volumes may reveal improved group discrimination of the AD subjects from normal controls. One issue which requires consideration is the problem of scale with regards to subdividing structures; a large structure will be viewed and segmented at a different level of magnification, and including or excluding some voxels in one part of the segmentation may not affect the fractional volume difference of that measure. However, if this large region is subdivided, the inclusion / exclusion of these voxels may be a relatively higher proportion of the volume of the structure which will have a greater effect on the fractional volume. Therefore a different level of attention to detail may be necessary in the volume estimation of the subregion. When subdividing a structure, it is important to assess whether it is appropriate in terms of accuracy and reliability or whether it is better to develop a strict protocol for that region alone.

2.12. Reducing amount of work in segmentation

2.12.1. Segmenting fewer slices

One way of reducing the amount of operator time when segmenting a region is to only measure a part of it, and estimate the volume of the missing slices by interpolation (see Figure 2.6). This method is particularly applicable when a structure is relatively smooth and areas between drawn slices do not deviate in size or shape such that interpolation

between slices when estimating volume is relatively accurate. If a structure is relatively similar throughout two planes e.g. a rectangle then there may be no benefit in measuring anything other than the first or last slice. In reality, there are few brain regions that meet these criteria and therefore this type of measure is not widely used when investigating a structure of interest. One example of a measure where this has been validated is TIV as the inside of the skull is usually a smooth structure (See Appendix Four for details and Figure 1.6) (Whitwell *et al.*, 2001).

2.12.2. Increased automation

A number of techniques have been described that may increase the degree of automation of delineation of structures in the brain. The need for accuracy of these segmentations depends upon the use of the mask, i.e. it may just be used for sampling voxel values within that region which may not require a high level of accuracy, or it may be the volume of that mask is required and therefore the need for accuracy is greater. Table 2.1 shows different studies which have used automated segmentation of the hippocampal region in order to assess cross sectional volumes in AD vs. controls or rates of change in those volumes in similar subject groups. Increased automation will be discussed in more detail later.

2.13. Inter-subject registration and correspondence

This chapter has only considered intra-subject registration until this point. However, increasing the automation of segmentation techniques often requires the registration of one subject's scan to another (often referred to as a template or atlas) on which segmentations of structures or tissue types have already been defined. Correspondence is most difficult to define when aligning scans of different subjects whether this is one subject to another or when combining many images to make a template image. Most inter-subject correspondences can then be defined either in anatomical terms, such as one temporal lobe aligning with all other temporal lobes, or in geometric terms, such as the matching of curvature of a gyrus between scans.

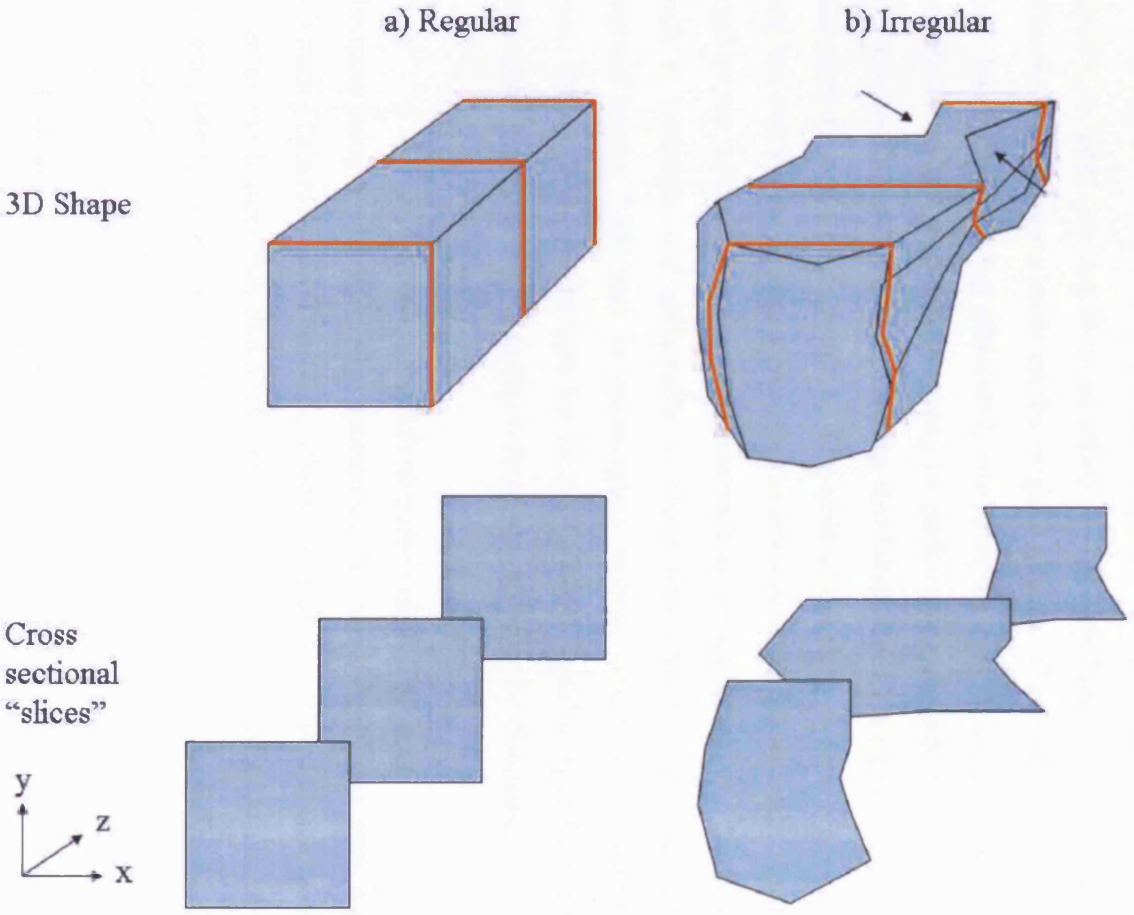


Figure 2.6 Reducing the number of slices in a segmentation.
Measuring fewer slices and interpolating the spaces between the segmented slices is accurate when a structure does not vary in terms of shape throughout the image (a). Inaccuracies in the estimation of volume by interpolating between the few measured slices will occur when regions vary in shape throughout the image particularly in areas between slices (b).

2.14. Templates: images and regions of interest

Templates are usually created to allow one of two things to happen; either to place many images into the same stereotaxic space for analysis (such as statistical parametric mapping (SPM) (Friston *et al.*, 1995b)) or VBM (Ashburner and Friston, 2000)) or to propagate labels in order to delineate specific structures (Fischl *et al.*, 2002) (see Figure 2.7). Templates are usually one of two types, either population (made from a number of subjects) or single-person based. Often these templates will have regions of interest delineated directly on them (making them an atlas) or will have labels which are averages of many different segmentations made on the original images. Choice of type of template depends upon the problem being addressed; this may be putting people into a similar spatial framework so that segmentations may be performed more precisely, so that differences in structure can be assessed, or it may be that labels from a template are to be transformed back onto the scans of individuals. The scale of the analysis also needs consideration as a template which may be appropriate for the analysis of the whole brain may be too sparse in detail to use with respect to small areas such as the hippocampus. Many of the advances made in templates are attributable to researchers involved in fMRI, where the signal detected by the fMRI may be meaningless without localisation to specific areas of the brain. Many of these templates for localising fMRI signal to specific brain structures are applicable to structural MRI and therefore these will also be discussed.

2.14.1. Single person and cohort templates

The most recognised and widely reported single person template or atlas was that made by Talairach (Talairach and Tournoux, 1988). This is, by definition, a space rather than an image as the images that make up this template are not contiguous in three dimensions. The template space was a framework developed from a 60 year old normally-aged female from France who came to post-mortem. Certain planes were defined in the template brain which form the basis of a co-ordinate system which can define tissue type according to position such that any given brain can be aligned and scaled to this space to define relevant areas. The main limitations of Talairach template space are that it is made up of one person, only one side of the brain was atlased and reflected and it may be subject to post-mortem artefact.

A cohort-based template is one where an additional subject is scanned within a study using the same imaging protocol and is then excluded from the group analysis but is instead used as a template image. The advantages to this type of template are that the scanner and protocol will be the same as those others in the study, reducing the amount of problems caused by such differences. Segmentations of any labels or regions on such images will be performed using the same protocol or interpretation of a protocol as those within a research group which reduces potential bias attributable to segmentation differences (Carmichael *et al.*, 2005). Limitations of this approach include the fact that these types of template do not allow for any of the normal variation seen between subjects; they may introduce bias as the template image is from one subject group and many subject groups are being assessed; they may cause difficulty in comparing results across studies; and results are subject to the quality of template image. However, for small structures, this may be the best form of template image to use as population-based templates may be too indistinct at smaller scales. A number of studies have utilised this form of template (Csernansky *et al.*, 2000; Wang *et al.*, 2003). Another template developed by Harvard is of a single patient who has many sections of the brain manually outlined, producing a well-defined atlas, (Kikinis *et al.*, 1996) and was originally designed for surgical planning, segmentation and teaching. The brain image and regions can be freely downloaded and as a result it has been used to segment the hippocampus in several studies (Carmichael *et al.*, 2005).

Some attempts have been made to circumvent the problems associated with single-person templates. One template developed by the Montreal Neurological Institute (MNI) was that of imaging the same subject 27 times. The average image of the registered images gave a very clear image with reduced noise. This template has been segmented to aid the localisation of signal on functional MR images in SPM (Tzourio-Mazoyer *et al.*, 2002). To limit the amount of bias introduced by having a single person template, another study has described a method of transforming the single-person template into the average space of all subjects in the study (Kochunov *et al.*, 2002). This technique not only reduces such bias but also retains the resolution and image quality associated with using a single person template.

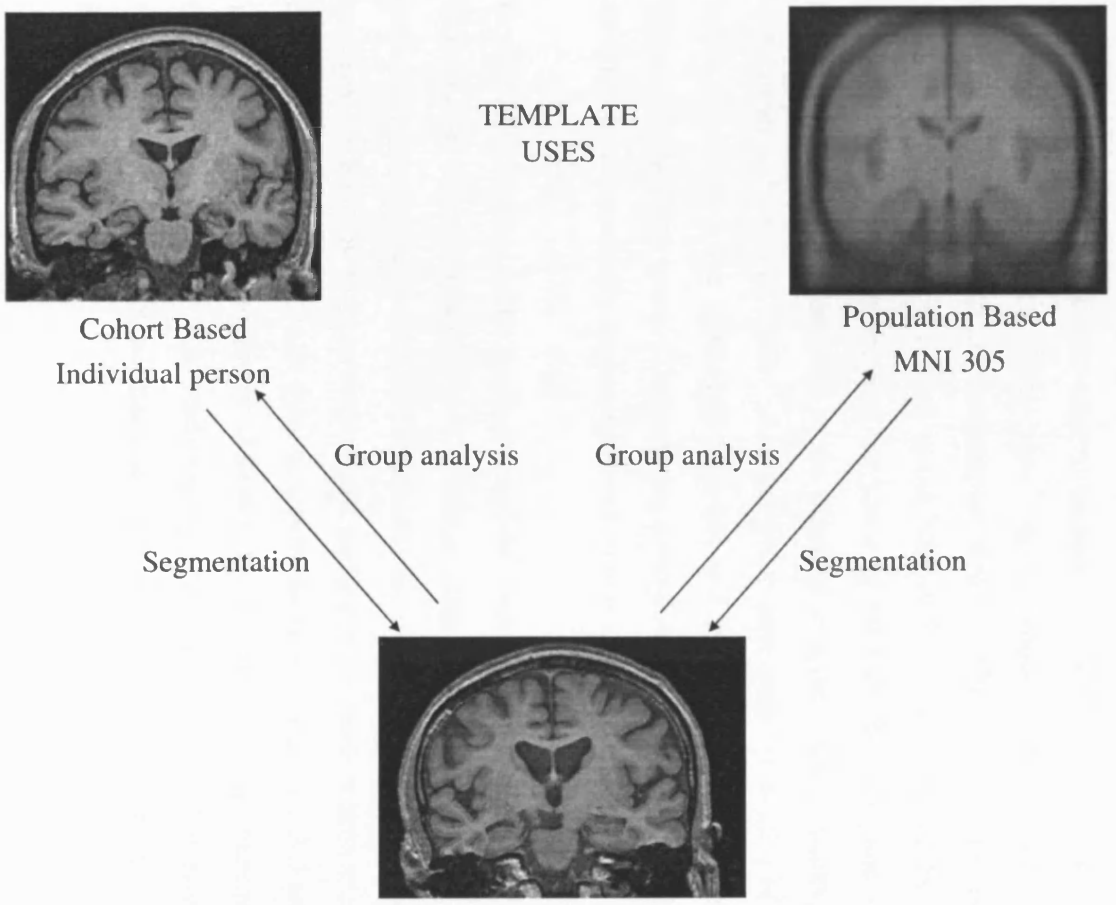


Figure 2.7 Schematic diagram showing basic uses of template images.
Arrows represent direction of registration.

2.14.2. *Population-based templates*

As a result of the problems associated with single-person templates, population-based templates have been developed. One example is the MNI305 template (Mazziotta *et al.*, 1991; Mazziotta *et al.*, 1995) (see Figure 2.7). The template was created from 305 normal control brains which were processed for radiofrequency inhomogeneity, aligned and averaged to create the image seen in Figure 2.7. This template image is in approximate Talairach space and functions exist to transform co-ordinates from the MNI 305 space to Talairach co-ordinates which are more widely used (see: http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml). One further template is named the ICBM152 template and is related to the MNI 305 as it was created by transforming 152 normal control subject scans into the space of the MNI 305 template and averaging the resultant images. This template is the standard template used in packages such as SPM (Ashburner and Friston, 1999). Another group have built up an atlas of young individual scans registered to the ICBM152 template and transformed the regions defined on the original scans to this space. The resulting average template image and template regions are designed to automatically segment regions on new scans (Hammers *et al.*, 2003).

In addition to normal-subject-group templates there is a need to develop templates that are disease specific. Thompson *et al.* have produced a template which is comprised of geometrically aligned scans of AD patients which creates an average map of the rendered surfaces. In addition to an average of the surface of the brain other variables such as grey matter density and variance in gyral structure can be mapped. The AD template developed was shown to be significantly different to that made using matched control scans (Thompson *et al.*, 2001). This study highlights the potential bias that may be introduced by using a template made of normal scans in a study where disease and controls are being assessed.

Although population-based templates may be useful in localising differences at a group level they have been shown to be less able to aid in the segmentation of small areas such as hippocampus compared with cohort templates (Carmichael *et al.*, 2005).

2.15. Automation of hippocampal measurements

Owing to early involvement of the hippocampus in AD and involvement of the structure in other diseases such as temporal lobe epilepsy and schizophrenia, efforts have been made to automate the segmentation of the structure to enable the volume to be estimated. In addition, the measurement of progression of atrophy in this area has also been targeted to improve upon entirely operator-dependent measures. A number of different techniques have been described to achieve these aims and these are described in Table 2.1. Many template images with labeled regions exist which were developed for alternative uses (for example fMRI activation map localisation (Hammers *et al.*, 2002; Hammers *et al.*, 2003; Tzourio-Mazoyer *et al.*, 2002), and although these could be used potentially for hippocampal segmentation or atrophy detection, these have not been included as they have not been formally assessed.

2.15.1. Cross-sectional hippocampal segmentation

The vast majority of segmentation techniques have been evaluated for cross-sectional hippocampal volume estimation. Comparisons between accuracy of different techniques are difficult to make owing to the different methods of assessing accuracy of segmentations. One paper compared a number of different registration techniques and templates but only freely available templates and registration techniques were considered (Carmichael *et al.*, 2005). Their main findings were that single-person cohort template images, site-specific hippocampal segmentations, and fully deformable registration methods improved hippocampal delineation. One automated technique has been applied in many further research studies (Csernansky *et al.*, 1998; Csernansky *et al.*, 2005; Wang *et al.*, 2001; Wang *et al.*, 2003). Not all techniques have been developed to segment hippocampi accurately; one technique samples the intensity within an approximate hippocampal region, made from the union of a number of manually-delineated control subjects (Webb *et al.*, 1999). Although this technique was applied to temporal lobe epilepsy patients it has not been applied to or assessed within AD.

2.15.2. Longitudinal hippocampal segmentation

Two techniques that claim to automatically segment hippocampal regions on serial images such that change in volume can be detected over time are those of (Du *et al.*, 2004; Wang *et al.*, 2003). These methods are based upon the application of the cross-sectional technique

described by Haller *et al.* (Haller *et al.*, 1997) applied to serial images. However, there is still some level of operator input required for this technique to achieve accurate results (manual placement of landmarks to constrict the warping of the template onto individual scans). One technique listed in Table 2.1 is semi-automated as it is necessary to manually delineate the baseline hippocampus in order for the region to be propagated by non-linear transformation derived from a local intra-subject fluid registration (Crum *et al.*, 2001). Although this method achieves highly accurate segmentations of repeat hippocampi, the limiting factor when considering the use of this technique is the need for the labour-intensive segmentation of the baseline region in order to generate a rate of atrophy.

Table 2.1 Semi-automated and automated hippocampal segmentation techniques

Study	**Objective of paper**	**Subject groups**	**Hippocampal volume/rate of change**	**Segmentation methods employed**	**Operator Input**	**Methods used to assess accuracy of segmentation compared with manual**
(Gosche *et al.*, 2001)	Assessing hippocampal volumes and relationship to Braak staging	51 PM confirmed cases. Mixed pathological statuses: different AD stages	Volume	Prior probability and intensity No template PM images only	Operator selects first section	Correlation (r) between 2 manual raters compared with correlation between each rater and the automated method
(Ghanei *et al.*, 1998)	Description of new method of segmentation	Only 2 hippocampi segmented. No relevant clinical information	Volume	Surface based deformable contours with edge tracking	Some manual intervention	Variation of similarity of voxel overlap on each slice of on hippocampus Comparison with manual segmentation
(Pitiot *et al.*, 2004)	Description of hippocampal segmentation method	Unclear	Volume	Mesh based registration and complex algorithm including shape and distance constraints	Unclear	Mean absolute surface difference of automated segmentation compared with manual

Study	Objective of paper	Subject groups	Hippocampal volume/rate of change	Segmentation methods employed	Operator Input	Methods used to assess accuracy of segmentation compared with manual
(Carmichael *et al.*, 2005)	Evaluation of publicly available registration algorithms and templates	20 AD, 19 MCI, 15 controls	Volume	Registration of atlas image to individual subjects and propagation of region	None	Similarity index
(Wang *et al.*, 2003)	Assessing volume and shape changes longitudinally	18 Mild DAT and 26 controls	Volume, rate and shape change	Haller *et al.* 1996-1997 technique used	Landmarks manually placed Cohort template delineated	None
(Webb *et al.*, 1999)	Automatic detection of hippocampal atrophy for use in epilepsy patients prior to surgery	15 TLEs 14 controls	Intensity within approximate hippocampal region	Union region made of 30 normal hippocampi, global and then local affine registration Analysis of intensity in ROI	None once union region made	None

Study	Objective of paper	Subject groups	Hippocampal volume/rate of change	Segmentation methods employed	Operator Input	Methods used to assess accuracy of segmentation compared with manual
(Shen *et al.*, 2002)	To asses the size and shape using a deformable shape model	10 subjects	Volume	Surface-based geometric model combined with prior probability map	Manual landmarks	Similarity index variation (overlap error as a percentage)
(Ashton *et al.*, 1997)	Assessing novel feature extraction technique in ability to assess asymmetry of volumes	9 TLE subjects	Volume	A priori shape and region growing	Manual planting of seeds	Percentage difference in volume. Slice by slice differences in volume.
(Crum *et al.*, 2001)	Assessment of fluid-based measures in detecting change over time	15 controls 12 AD	Rates	Intra subject fluid Manual (outline baseline)	Segmentation of baseline hippocampus	Voxel difference using same day scan-pairs
(Haller *et al.*, 1997)	Assessing template based segmentation methods	9 controls 9 schizophrenic patients	Volumes	Non-linear (fluid) registration of template hippocampus onto new scans.	Segmentation of the template hippocampus Manual landmarks	Percentage difference in volume of the automated technique performed twice. Percentage difference in volume of the manual performed twice

Study	Objective of paper	Subject groups	Hippocampal volume/rate of change	Segmentation methods employed	Operator Input	Methods used to assess accuracy of segmentation compared with manual
(Duchesne *et al.*, 2002)	Evaluating new method of segmentation against manual and existing automated method	80 normal subjects	Volumes	Intensity based methods with shape description	Manual segmentation of training set	Similarity index compared with manual segmentation
(Fischl *et al.*, 2002)	Description of new method of propagating many labels from a template	134 subjects: 25 controls, 92 AD with various longitudinal outcomes	Volumes	Registration of a template to subjects with incorporation of spatial and probabilistic information	None	Version of similarity index.
(Hsu *et al.*, 2002)	Assessing template based segmentation methods	20 controls, 20 cognitively impaired and 20 AD subjects	Volumes	Technique based on Haller *et al.* 1997, i.e. segmentation of template hippocampus followed by fluid-based registration and segmentation of hippocampi constrained by manually-placed landmarks	Segmentation of the template hippocampus and placement of manual landmarks	ICC of automated volumes compared with ICC of manual volumes. These measures based on 10 elderly subjects.

Study	Objective of paper	Subject groups	Hippocampal volume/rate of change	Segmentation methods employed	Operator Input	Methods used to assess accuracy of segmentation compared with manual
(Du *et al.*, 2004)	Establishing whether hippocampus or entorhinal cortex atrophy rate is higher	25 Normal controls 20 AD subjects	Rate of change	Same as Hsu *et al.* 2002 performed at two time-points per subject allowing rate of change to be calculated	Segmentation of the template hippocampus and placement of manual landmarks	None

Key

TLE temporal lobe epilepsy
DAT dementia of Alzheimer type
PM Post-mortem
AD Alzheimer's disease
MCI Mild cognitive impairment

2.16. Chapter conclusions

Both segmentation and registration have been discussed in this chapter as these two parts of image analysis are often intertwined; accurate automated segmentation is difficult to perform without registration and accurate automated registration often requires some form of prior segmentation to limit the areas over which the registrations are performed. Segmentation and registration form the backbone of many image analysis techniques. However, there is still much work to be done in many areas; either to either improve accuracy of automatic segmentations or measure rates of atrophy or to allow analysis of many subjects within a true stereotaxic spatial framework where accurate correspondences have been achieved.

3. METHODS OVERVIEW

3.1. Subjects

Subjects for the studies described in this thesis were recruited from the following sources:

3.1.1. The Specialist Cognitive Disorders Clinic

This clinic is based at the National Hospital for Neurology and Neurosurgery, Queen Square, London. This is a tertiary referral centre, and consequently these subjects tend to represent younger patients and those where there is more diagnostic uncertainty than a typical elderly population. The subjects recruited from the clinic included individuals with sporadic or familial AD, FTLD and other more unusual dementias. All subjects underwent full clinical and neuropsychological assessment. Diagnosis was made according to the criteria detailed in Appendix One.

3.1.2. Familial AD project

This is an ongoing longitudinal research project following individuals with a strong family history of histologically-proven AD. These families have, in many cases, subsequently been found to have mutations in the PS1 or APP genes. Affected individuals and asymptomatic family members at risk of developing the disease undergo annual assessments. These include a full history and neurological examination, MMSE, detailed neuropsychology and an MRI scan. This provides valuable information about presymptomatic individuals who subsequently go on to develop the disease, aiding in the understanding of the earliest signs of the disease. This cohort currently includes 30 affected and 48 at-risk individuals under follow-up with clinical and imaging assessments.

3.1.3. MIRIAD

The Minimum Interval Resonance Imaging in AD (MIRIAD) project aimed to determine the shortest interval required to detect volumetric change based on MRI, to distinguish AD subjects from normal controls. Approximately 50 individuals with sporadic AD and 26 control subjects were recruited and followed over a period of at least one year. These subjects mainly consisted of individuals over the age of 55 attending the Specialist Cognitive Disorders Clinic and their spouses. Individuals with AD had MMSE scores in the range of 12-26 at baseline assessment.

Subjects were assessed at baseline and at 2, 6, 14, 26, 38 and 52 weeks. At baseline all subjects underwent a full history, neurological examination, MMSE and detailed neuropsychology. Speed tests were performed at each visit and MMSE was repeated at 26 and 52 weeks. At 52 weeks the initial clinical diagnosis was reassessed and neuropsychology was repeated. Apolipoprotein (APOE) genotyping and electroencephalograms (EEG) were performed on patients alone. A volumetric MR scan was performed at each visit. At baseline, 6 weeks and 38 weeks two scans were acquired on the same day. In addition, MRS was performed at weeks 2, 26 and 52. A subset of these subjects (n=39) were assessed and scanned at 18 months, and at two years (n=22).

3.1.4. Normal controls

Neurologically healthy individuals with no family history of dementia are recruited to act as normal controls. These are largely recruited from spouses of affected or at-risk study members and a number of normal volunteers. All subjects have a detailed history taken and undergo a neurological examination, MMSE and a brief memory test, as well as MRI assessment. The spouses of study subjects are usually scanned on the same day as their partners and therefore have as closely comparable MR acquisitions as possible.

3.1.5. Pathologically-proven cases

As a result of being a patient at the Specialised Cognitive Disorders Clinic, a number of patients undergo brain biopsy to determine the cause of disease (see section 3.2). In addition to this, a number of patients followed through clinic have agreed to a post-mortem to accurately diagnose the cause of dementia. Such cases are valuable as pathological evidence of disease is the "gold standard" of diagnosis. As a result, the scans obtained for clinical reasons, prior to death are a valuable resource for assessing group separation using different methodologies as there is diagnostic certainty. Some of these subjects were enrolled in other studies, such as MIRIAD. Seventy-nine subjects have had post-mortem confirmation of disease and at least one scan. Approximately half of these subjects have had a number of scans although these scans are of variable quality and some longitudinal imaging was performed on different scanners.

All subjects gave written informed consent for involvement in imaging studies and the Local Ethics Research Committee had given approval for all studies.

3.2. Clinical assessment

All subjects attending the Specialist Cognitive Disorders Clinic undergo comprehensive diagnostic evaluation. A full history, including family history, is taken together with a close informant. A full medical examination is performed as well as a neurological assessment. In addition, almost without exception, the following investigations are usually performed:

1. Detailed neuropsychology, in order to establish the nature and severity of any cognitive deficits.
2. Standard screening blood tests to exclude other treatable causes of cognitive problems such as impaired renal or liver function, B12 and thyroid function.
3. EEG to exclude seizures, or identify patterns indicative of a particular type of dementia.
4. Neuroimaging will typically involve at least one MRI assessment. This can exclude treatable causes such as tumours and subdural haematomas, or may indicate progressive neurodegeneration suggestive of some type of dementia pathology.

In the appropriate setting individuals may also undergo:

1. Genetic testing. In the case of individuals with a known genetic mutation in the family, genetic testing may be offered to determine whether a subject is carrying the mutation. This is supported by a full genetic counselling service.
2. Lumbar puncture and CSF analysis. This investigation can be used to detect inflammation due to e.g. meningitis or encephalitis, or to detect local CNS immune responses due to e.g. MS.

In a small number of subjects a brain biopsy may be performed. This is an invasive procedure which involves taking a small amount of brain tissue usually from the right (or non-dominant) frontal lobe. Definitive diagnosis can be obtained in this way by identifying a particular pathology. This technique is not used routinely, but may be suggested in the case of young individuals with dementia where it may be more likely that there are other treatable causes, especially if cerebral vasculitis is considered.

3.3. *Imaging*

3.3.1. *Acquisition*

All images were acquired on a 1.5T Signa MRI scanner (General Electric, Milwaukee, Wisconsin, USA) using a spoiled gradient echo technique. Scans generally included a sagittal T1-weighted scout sequence, an axial dual-echo sequence (T2-weighted and proton-density weighted) and a T1-weighted volumetric image (124 contiguous 1.5mm slices). Imaging parameters varied according to the study, and details are given in Appendix Three.

3.3.2. *Software and processing*

Digitised images were transferred to a Sun workstation (Sun Microsystems Inc., Mountain View, CA) for analysis. Images which were judged on visual inspection to show a significant intensity gradient were corrected using the N3 correction algorithm (Sled *et al.*, 1998) outlined in Chapter 2 (page 71).

3.3.2.1.*MIDAS*

The MIDAS (Medical Information Display and Analysis System) software (Freeborough *et al.*, 1996a; Freeborough *et al.*, 1997) runs on both Unix and Linux platforms and is implemented in the C programming language. This software allows simultaneous multiplanar display of 3D data. Brain structures can be outlined using both semi-automated and manual techniques. The whole brain segmentation tool within MIDAS is semi-automated and uses interactive thresholding, together with a series of erosions and dilations, to isolate brain tissue from other structures such as scalp and dura (see Figure 1.6 and Appendix Four).

Manual segmentation of brain sub-structures, such as the hippocampus, can be performed using a mouse-driven cursor. The simultaneous display of orthogonal views allows the operator to outline the structure in the coronal view whilst the segmentation is updated in real time in the sagittal or axial view. This aids in decisions about where boundaries should be defined (see Appendix Four). Both rigid-body, rigid-body plus scalings and shears, and fluid registration, as described in Chapter 2, can be implemented within the MIDAS software.

3.3.2.2.STATA

STATA (Stata Coporation, College Station, TX, USA) is a standard statistical package used for producing graphs and performing basic statistical analysis. This was used for the majority of statistical analysis and graph production within this thesis.

3.3.2.3.SAS

SAS (SAS Institute, Inc, Cary, NC, USA) was used for some statistical analysis as it allows complex models to be created where variances and covariances do not have to be similar.

4. CROSS SECTIONAL RESULTS USING MANUAL SEGMENTATION

4.1. Chapter introduction

The "gold standard" of region delineation

Manual segmentation of regions within the brain is currently the "gold standard" of volumetric analysis. Although some semi-automated techniques exist which enable delineation of substructures (see Table 2.1 for examples of hippocampal segmentation), none of these have become completely automated and been clinically validated against manual measures. Assuming scanning protocols and qualities remain constant across a study, cross-sectional measures give a quantification of the volume of a structure which represents i) the amount of atrophy which has occurred up to the point at which the scan was taken, ii) the normal inter-individual variability in structural volume iii) errors in delineation caused by incorrect interpretation of protocols and iv) noise in the measurements made by the scanner. The first of these measures of variability can be caused by disease and result in some disease groups having smaller regions in specific areas of the brain compared with other neurological diseases or control groups. As a result, researchers are interested in i), taking into account ii) whilst reducing or eliminating iii) and iv). Some aspects of the inter-individual variability mentioned in ii) can be reduced by head-size correction such as TIV adjustment, however if there is an association between TIV and disease, this can reduce sensitivity of the measures. The third source of variability (manual delineation errors) can be controlled to some extent by repeated measurement and by time consuming careful editing. This chapter concentrates on the resulting "gold standard" cross-sectional volumetric measures.

Structures delineated in AD and FTLD

The subregions usually targeted for delineation are often those for which there is existing evidence of the region being affected in some way by the disease process. This evidence is most often demonstrated by either visual inspection of MR images, which often report reduction of specific areas of the brain (Likeman *et al.*, 2005), or differences noticed between disease groups and controls within another imaging modality such as MRS (Kizu *et al.*, 2004), SPECT (Matsuda *et al.*, 2002) or PET (Valla *et al.*, 2001). Other sources of evidence include pathological studies, some of which describe stages of disease and therefore give potential targets for early diagnosis (Braak *et al.*, 1993). Within Alzheimer's disease there are many regions which have been assessed; these are usually within the temporal lobe or are part of the limbic system as

these areas are known to be affected by the disease, with specific structures affected early in the disease course (entorhinal cortex, hippocampus, amygdala and to a lesser extent the cingulate gyrus). In FTLD, studies on temporal lobe structures have also been performed, but owing to the heterogeneity of the clinical syndromes and the location of atrophy, more anterior regions such as frontal lobes have also been assessed. Results from some cross-sectional volumetric studies in AD are summarised in Table 1.1.

Limitations

There are limitations to the application or the interpretations of the results of manually-performed volumetric cross-sectional techniques. One is that developing reliable protocols can take a considerable length of time; even when such protocols have been developed they can take much time to apply to new scans and as a result these techniques are not applicable in routine clinical situations. With such studies, the division of two types of variation (inter-individual and prior atrophy) cannot be differentiated; for this serial analysis of scans is required. Although differences in sizes of structures have been found between many different disease groups, the diagnostic utility of such measures is limited by the natural variation in sizes of structures even following whole head size correction.

Chapter outline

In this chapter the application of new protocols to delineate structures of interest are described and assessed. In addition, existing protocols for the hippocampus and amygdala are applied to groups of pathologically-confirmed subjects.

4.2. Developing protocol for ROI delineation for cross sectional studies

4.2.1. Introduction

MRI-based volumetric measurements have proven to be useful for quantifying regional cerebral atrophy occurring in AD (Fox *et al.*, 2001). Post-mortem studies have demonstrated that the atrophy in AD is associated with its defining histopathological changes, i.e. the accumulation of neurofibrillary tangles and amyloid plaques, accompanied by widespread neuronal and synaptic loss (Braak and Braak, 1991; Delacourte *et al.*, 1999; Gomez-Isla *et al.*, 1996; Rohn *et al.*, 2001; Uylings and de Brabander, 2002). The implication therefore, is that the macroscopic changes of

progressive regional and global atrophy are a closely linked consequence of the underlying pathological processes in AD.

As discussed earlier in Chapter 1 (see page 41), the hippocampus and entorhinal cortex are thought to be the earliest sites of pathological involvement in AD at least by neurofibrillary tangles (Braak and Braak, 1991) and many volumetric MRI studies have focused on these regions in sporadic and familial AD, confirming their early involvement *in vivo* (Fox *et al.*, 1996a; Jack *et al.*, 1999; Schott *et al.*, 2003). More recently, volumetric, functional and neuropathological studies have also highlighted another area of the brain that appears to be involved from the very earliest stages of the disease: the cingulate cortex (Baron *et al.*, 2001; Braak and Braak, 1993; Callen *et al.*, 2001; Fox *et al.*, 2001; Frisoni *et al.*, 2002a; Killiany *et al.*, 2000; Minoshima *et al.*, 1997; Scahill *et al.*, 2002).

The cingulate cortex is a structurally and functionally heterogeneous region, located on the medial surface of the brain. It can be subdivided into an anterior cingulate (AC) and a posterior cingulate (PC) cortex (Baleydier and Mauguiere, 1980; Vogt *et al.*, 1979). The PC is subdivided further into the ventro-medially located retrosplenial cortex (RS) and the more dorsally located PC 'proper', (Vogt *et al.*, 2001). The RS has particularly dense connections with the medial temporal lobe and demonstrates neurofibrillary changes at an earlier histopathological stage of AD, compared with the rest of the PCC (Braak and Braak, 1993; Insausti *et al.*, 1987; Kobayashi and Amaral, 2003; Lavenex *et al.*, 2002; Suzuki and Amaral, 1994). In view of these connectional and neuropathological differences, cingulate subregions may differ with respect to their amount and rate of atrophy in AD. Methods to assess regional cingulate atrophy on MRI accurately, may therefore provide insights into how the disease progresses and might ultimately be used to improve the diagnostic accuracy in very early stages of the disease (Baron *et al.*, 2001; Chetelat and Baron, 2003; Kelly *et al.*, 1997).

Previous imaging studies have demonstrated atrophy in the posterior cingulate gyrus in AD. However, these studies either used different criteria to delineate and subdivide the cingulate cortex and did not subdivide the PC (Callen *et al.*, 2001; Killiany *et al.*, 2000), or used semiautomatic techniques (Baron *et al.*, 2001; Fox *et al.*, 2001; Frisoni *et al.*, 2002a; Scahill *et al.*, 2002), which may lack anatomic specificity for relatively small structures such as the RS (Ashburner and Friston, 2000; Crum *et al.*, 2003; Tisserand *et*

al., 2002). Several studies have demonstrated atrophy in the AC in later stages of the disease but this was variably located in its rostral (Frisoni *et al.*, 2002a; Scahill *et al.*, 2002) or its more caudal parts (Callen *et al.*, 2001; Killiany *et al.*, 2000).

The aim of this study was to develop and validate a delineation protocol, suitable for measuring atrophy of the different regions of the cingulate gyrus and subsequently to examine the regional pattern of cingulate atrophy in AD. For this purpose the volumes of the different cingulate regions of interest in early-onset familial AD patients were compared with those of age and sex-matched control subjects. Early-onset FAD patients were studied as the diagnosis of AD can be made with relative certainty owing to their known genetic risk. Moreover, in these patients co-morbidity, such as vascular disease, is less of a confound. It was also important to assess whether this protocol could reproducibly detect cortical volume reduction in smaller regions of the cingulate gyrus such as the retrosplenial cortex.

4.2.2. *Subjects and methods*

Ten subjects with familial Alzheimer's disease, fulfilling criteria for probable AD (McKhann *et al.*, 1984) (see Appendix One) and ten healthy age- and sex-matched controls recruited from the spouses of patients and healthy volunteers were selected for this study. Eight patients had genetic testing that confirmed mutations known to be pathogenic for AD (six in APP; two in PS-1) and two patients came from autosomal dominant pedigrees with pathology in a first degree relative. Subject demographics are described in Table 4.1. Subjects underwent annual MRI, MMSE (Folstein *et al.*, 1975) (see Appendix Two) and detailed clinical and neuropsychological assessment as part of an ongoing longitudinal research project. Subjects had no medical history of cerebrovascular or other chronic neurological disease, systemic disorders or major psychiatric illnesses. The study was approved by the Local Research Ethics Committee and all subjects had given written informed consent.

Table 4.1 Subject demographics

	Controls	AD subjects
Number	10	10
M/F	5/5	4/6
Mean age, years (SD)	51.0 (8.0)	51.8 (7.4)
MMSE score /30		
mean (SD)	29.9 (0.4)	10.6 (6.2)
range	29-30	5-21

4.2.2.1.Magnetic resonance imaging acquisition

T1-weighted volumetric MR brain scans were acquired as described by the standard protocol in Appendix Three. The scans were acquired as 124 contiguous 1.5mm coronal slices and were transferred to a Sun workstation (Sun Microsystems Inc., Mountain View, CA) for analysis.

4.2.2.2.Image analysis

Prior to segmentation, all scans were globally registered to the MNI 305 template using a six degrees of freedom registration algorithm (Mazziotta *et al.*, 1995). This ensured all scans were measured in a similar orientation to improve consistency of segmentation. The software package MIDAS was used for all manual segmentation (see Chapter 3, page 95). In order to prevent potential laterality bias, each image was presented twice in random order; once normally and once flipped across the midsagittal line, producing two scans, each a mirror image of the other. This enabled the ROIs to be consistently measured on the right-hand side of the presented image, whether the ROI was left or right. This methodology is hereafter referred to as "mirror-image volumetry". Measurements were performed by raters who were blind to the clinical diagnosis.

Cingulate borders were established using a combination of intensity thresholding and manual tracing using a mouse-driven cursor. A threshold setting between 70-110% of the mean brain intensity of the whole brain was used to give consistent delineation of cortical grey matter from CSF and white matter. Whole brain segmentation was performed using a previously described technique (Freeborough *et al.*, 1997; Schott *et al.*, 2003). All regional cingulate boundaries were manually outlined and edited.

4.2.2.3.Reproducibility

In order to assess intra- and inter-rater reproducibility, regional cingulate volumes of 10 patients (20 hemispheres) were each measured twice by two different raters. Subdivisions of the posterior cingulate gyrus (PC, RS) were measured twice by one rater, in order to assess the feasibility of separating these two regions in the further analysis. Reproducibility for the whole cingulate gyrus and component parts was assessed by calculating within-subject standard deviations and reliability coefficients (Bartko, 1966; Ebel, 1951; Fleis, 1986).

4.2.2.4.Statistical analysis

Statistical analyses were conducted using the statistical package STATA versions 6 and 8 and Microsoft Excel 2000. In order to normalize scans for individual differences in head size, TIV was calculated from T1-weighted scans, using a previously described semi-automated technique (Whitwell *et al.*, 2001) (see Appendix Four). The logs of the mean cingulate ROI volumes of the control group were regressed against the logs of the mean TIV measures to establish the slope of the relationship between TIV and cingulate volume (Free *et al.*, 1995). The resulting coefficient (α) was used to correct the cingulate volumes as follows:

$$h_i^* = h_i \, (t_0 / t_i)^{\alpha}$$

Equation 4.1

where h_i^* is the adjusted cingulate volume, h_i the crude cingulate volume, t_i the TIV for the *i*th individual and t_0 the mean TIV.

Differences in volumes between a) AD-cases and controls, b) cingulate subdivisions and c) left and right sides were assessed using a random effects generalised least squares regression model. All volumes were log-transformed (to allow estimation of percentage differences) and two-way interaction terms were included to investigate the extent to which a) the difference between AD-cases and controls differed between regions, b) the left-right difference differed between regions and c) the left-right difference differed between AD-cases and controls. Pair-wise contrasts were used to investigate interactions where a global interaction test was statistically significant.

4.2.3. The protocol

4.2.3.1.Cingulate borders

The cingulate gyrus largely comprises Brodmann's areas 24, 23, 29, 30 and part of 31 and effort was taken to confine the analysis to these regions as much as possible. The rules that were applied for delineation of cingulate boundaries and subdivisions were based upon cytoarchitectonic, connectional and macroscopic studies (Groenewegen and Uylings, 2000; Ono *et al.*, 1990; Sanz Arigita *et al.*, 2003; Tisserand *et al.*, 2002; Uylings *et al.*, 2000; Vogt *et al.*, 1995; Vogt *et al.*, 2001; Vogt and Vogt, 2003). Indicated below is a step-by-step procedure which describes how the macroscopic MRI delineations of the ROI can be roughly compared with microscopically defined Brodmann areas. It is essential to emphasise explicitly, that macroscopic features like sulci often do not coincide precisely with cytoarchitectonic borders (Uylings *et al.*, 2005; Zilles, 2004).

4.2.3.2.Sulcal patterns and variability: Demarcation of cingulate boundaries

Figure 4.1 illustrates the gyri and sulci used to demarcate cingulate gyrus borders. The cingulate gyrus (CG) arches around the corpus callosum on the medial surface of the brain and is separated from this structure by the callosal sulcus (*cas*) (Figure 4.1.A). The cingulate sulcus (*cs*) forms the border between the cingulate gyrus and the antero-dorsally located paracingulate gyrus (PCG) (Paus *et al.*, 1996; Tisserand *et al.*, 2002) and the (pre)motor cortex (MC) (Figure 4.1.A). The paracingulate gyrus may be heavily segmented (Figure 4.1.A) or more continuous (Figure 4.1.B) and may share a common origin with the cingulate gyrus (Figure 4.1.C). Rostral parts of the cingulate gyrus often contain a shallow secondary sulcus, running parallel to the cingulate sulcus (Ono *et al.*, 1990) (Figure 4.1.A-B, white arrowheads). Even in these cases the cingulate gyrus was

always easy to distinguish from the antero-dorsally located paracingulate gyrus. The cingulate sulcus continues up until the appearance of the marginal ramus (*mr*), from whence it continues as the splenial sulcus (*sps*) (Vogt *et al.*, 1995) (this has also been termed subparietal sulcus (Ono *et al.*, 1990)) (Figure 4.1.A). The splenial sulcus is highly variable in its pattern and often contains extensions directed towards the corpus callosum, which were included in the analysis (Figure 4.1.B-C, black arrowheads). The cingulate sulcus and splenial sulcus were sometimes interrupted and in such cases the shortest possible line between the interrupted segments of the sulcus was drawn (see for example Figure 4.1.C, curved arrow).

As the cingulate gyrus curves around the anterior part of the corpus callosum a transition takes place from Brodmann's area 24 to 25. In accordance with Tisserand (Tisserand *et al.*, 2002), the most caudal coronal slice on which the inner curvature of the genu of corpus callosum was visible was taken as the ventral-posterior border between areas 24 and 25 (Figure 4.1.B). In this manner area 25 was approximately excluded from the analysis. Area 25 is a more simple cortical structure than 24 (Vogt *et al.*, 1995; Vogt *et al.*, 2004). Posteriorly, the cingulate gyrus curves around the splenium of the corpus callosum where it ventrally borders the parahippocampal cortex. The most ventral axial slice on which the curvature of the splenium of the corpus callosum was visible has been taken as the macroscopic ventral border of the cingulate gyrus with the parahippocampal cortex, since no cingulate cortex is present below this edge (Vogt *et al.*, 2001) (Figure 4.1.B).

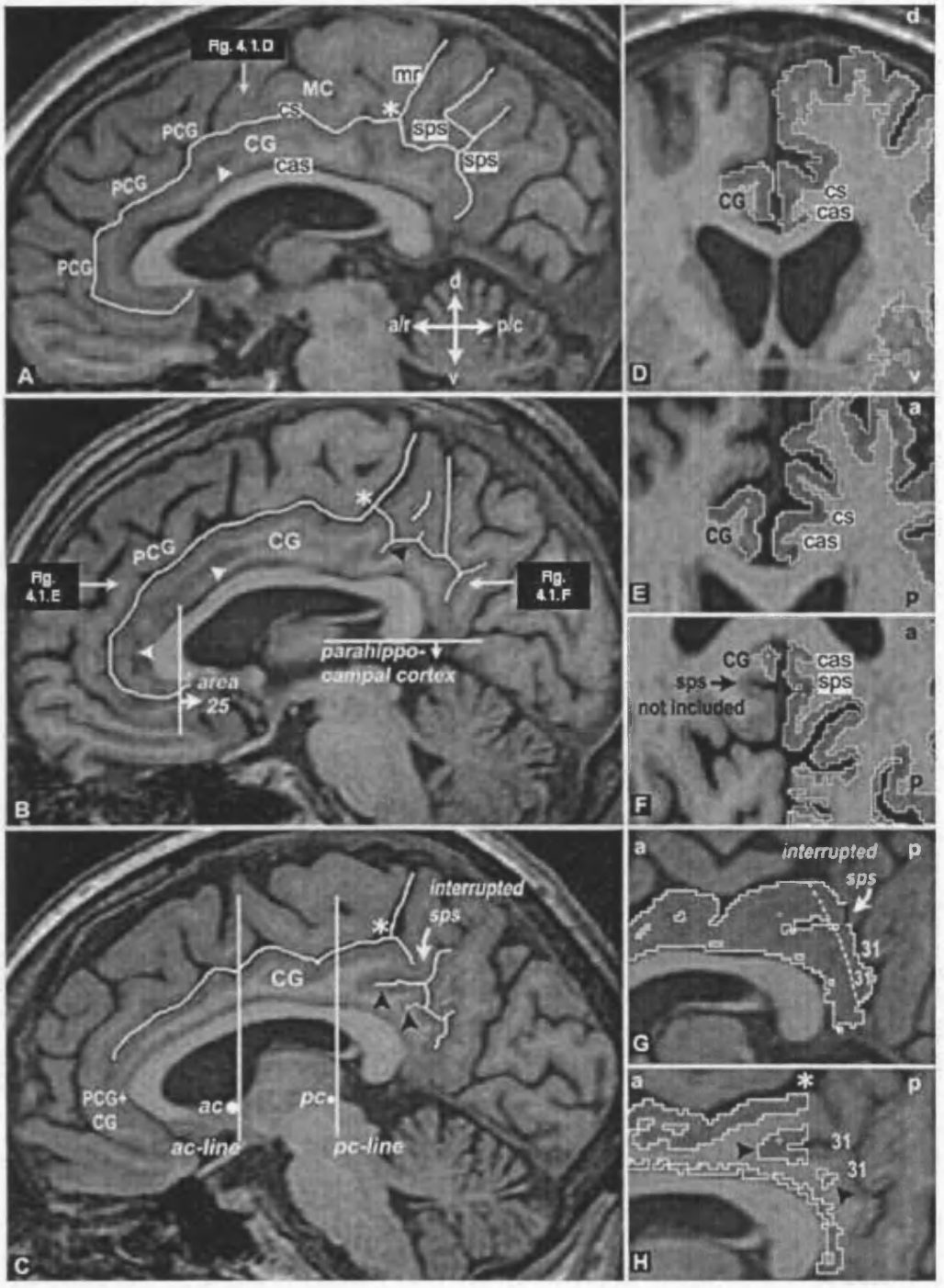


Figure 4.1 Demarcation of the cingulate gyrus.
(A-C) Sagittal sections of MRI scans of 3 representative brains indicating the sulci, gyri, and reference points used to demarcate cingulate gyrus borders. (D) Coronal, (E,F) axial, and (G, H) sagittal sections, demonstrating cingulate gyrus borders within the relevant sulci. The asterisk indicates the site where the cingulate sulcus, splenial sulcus, and marginal ramus meet. Ac-line and pc-line in C: cutoff sections through the posterior edge of the anterior commissure (ac) and the posterior edge of the posterior commissure (pc) respectively, used for subdividing the cingulate gyrus. a/r = anterior/rostral; cas = callosal sulcus; CG = cingulate gyrus; cs = cingulate sulcus; d = dorsal; MC = (pre)motor cortex; mr = marginal ramus; p/c = posterior/caudal; PCG = paracingulate sulcus; sps = splenial sulcus; v = ventral: 31 = Brodmann area 31.

Cingulate areas located dorsal to the corpus callosum were delineated on coronal slices (Figure 4.1.D), whereas cingulate areas located anterior and posterior to the corpus callosum were delineated on axial slices (Figure 4.1.E, F) respectively. On the right side of Figure 4.1.D-F the cortical grey matter has been automatically delineated from CSF and white matter using a threshold setting between 70-110% of the mean brain intensity of the whole brain. On the left side of the figures the cingulate gyrus has been delineated. Cingulate gyrus borders were always set between the ventral and dorsal banks of the callosal sulcus and the cingulate sulcus (Sanz Arigita *et al.*, 2003). The grey matter of the splenial sulcus was not included in the analysis (Figure 4.1.F). The splenial sulcus is surrounded by Brodmann's area 31 (Vogt *et al.*, 2001). Effort was taken to exclude the major part of area 31, since it is also located on the parasplenial lobules postero-dorsal to the cingulate gyrus (Vogt *et al.*, 2001), which was not a region of interest (illustrated on sagittal sections in Figure 4.1.G-H). However, a small part of area 31 is located on the medial surface of the cingulate gyrus and cannot be delineated macroscopically (Figure 4.1.G, approximate border indicated by dotted line). Figure 4.1.G is a more medial section of the brain in Figure 4.1.C. In order to improve consistency the entire medial surface of the cingulate gyrus was included in the analysis and thus also included a small part of area 31. The grey matter located within the splenial sulcus was removed in sagittal slices (Figure 4.1.H). Note that in Figure 4.1.H extensions of the splenial sulcus directed towards the corpus callosum were included in the analysis (black arrowheads).

4.2.3.3.Cingulate ROI subdivisions

The most caudal coronal slice on which the anterior commisure (*ac*) was visible (Figure 4.1.C, *ac-line*) and the most caudal coronal slice on which the posterior commisure was visible (Figure 4.1.C, *pc-line*) were used to subdivide the cingulate gyrus. The orientation of the coronal slices in the scans corresponds with the *ac-* and *pc-line* in Figure 4.1.C. The resulting subdivision of the cingulate gyrus and its borders within the sulci, together with the approximate location of the relevant Brodmann's areas, is indicated in Figure 4.2. Anterior parts of the cingulate gyrus were subdivided into a rostral ROI (termed rostral AC (RAC), located anterior to the *ac-line*) and a caudal ROI (termed caudal AC (CAC), located between the *ac-* and *pc-line*) (Figures. 4.2.A-C). Although both anterior cingulate regions approximately correspond to Brodmann's area 24, more recent studies have demonstrated a number of cytological and functional differences between them justifying a subdivision (Vogt *et al.*, 2003; Vogt and Vogt,

2003). Although the cingulate and paracingulate gyrus share a common origin in this specific brain a slight indentation is visible in the medial part of the grey matter, which was taken as the border between these regions (Figure 4.2.B, arrow). The posterior region of the cingulate gyrus (located caudal to the *pc-line*) was divided into a posterior cingulate ROI (termed PC) and the retrosplenial cortex (RS) (Figure 4.2.A, D, E). The border between the PC and the RS is the only border that can be delineated objectively and macroscopically (Vogt *et al.*, 2001). The retrosplenial cortex corresponds with Brodmann's areas 29 and 30 and is located in the depth of the callosal sulcus and does not appear on the medial surface of the cingulate gyrus, which comprises area 23 (Figure 4.2.D-E). Thus, when manually indicating the border of the RS (black line in Figure 4.2.D-E) care was taken not to include any medially located voxels.

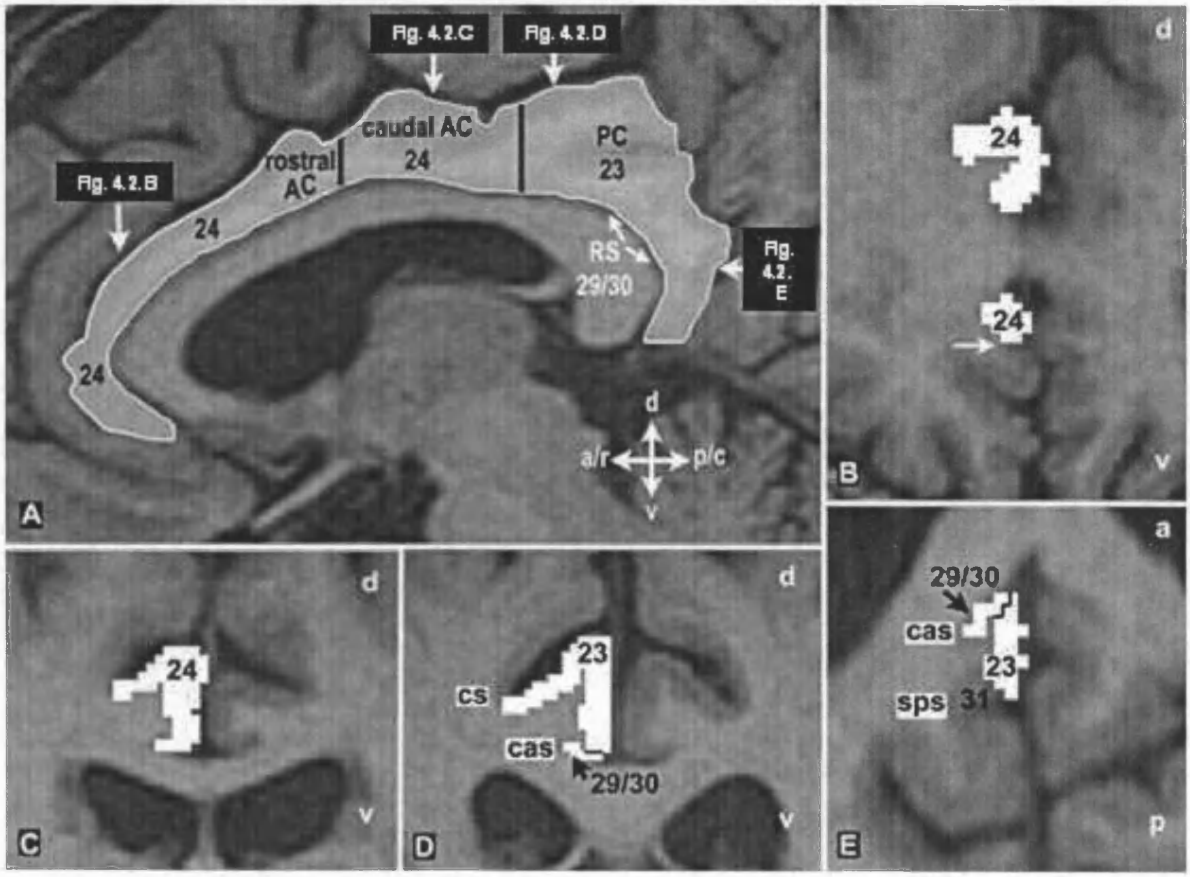


Figure 4.2 Subdivisions of the cingulate gyrus.
(A) Sagittal, (B--D) coronal, and (E) axial views of the same MRI scan illustrating the location of the cingulate subdivisions together with approximate corresponding Brodmann areas. (A) The orientation and location of the sections in 2B--E have been indicated. (B, C) Coronal sections through the rostral and caudal AC respectively. (D) Coronal and (E) axial section through the PC and RS. Abbreviations are the same as in Figure 4.1.

4.2.4. *Results*

4.2.4.1.*Reproducibility*

Reproducibility, expressed in terms of reliability coefficient and SDs as percentages are given in Table 4.2. Intra-and inter-rater reliability coefficients were comparable and ranged from 91.4%-99.4%. Considering the high reproducibility for the PC and the RS (>91.9%, (Donner and Eliasziw, 1987)), these two regions were fitted separately into the statistical model and not pooled together.

Table 4.2 Reproducibility of cingulate segmentation.

	Intra-rater reproducibility				***Inter-rater reproducibility***			
	Left		**Right**		**Left**		**Right**	
	RC(%)	**SD(%)**	**RC(%)**	**SD(%)**	**RC(%)**	**SD(%)**	**RC(%)**	**SD(%)**
Cingulate	99.4	2.4	98.5	2.8	97.9	4.1	99.1	2.2
Rostral AC	97.8	5.6	99.1	2.7	93.5	9.3	99.2	2.5
Caudal AC	97.9	3.9	97.8	3.8	98.8	2.8	98.4	3.2
PC + RS	96.2	7.9	98.5	4.4	94.9	9.2	98.9	3.8
PC	94.4	10.4	98.3	5.1	*	*	*	*
RS	91.9	9.1	92.4	8.3	*	*	*	*

RC = reliability co-efficient, AC = anterior cingulate, PC = posterior cingulate, RS is retrosplenial cortex

4.2.4.2.Cingulate subdivision volumes

The TIV corrected cingulate ROI volumes for both groups are shown in Figure 4.3.

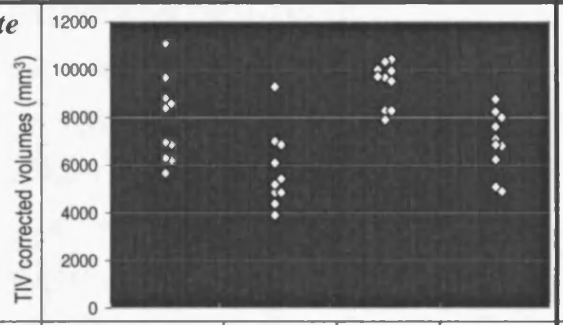
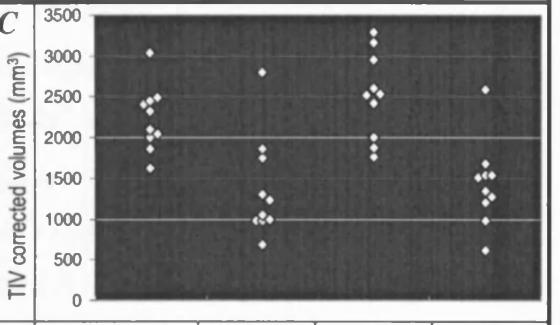
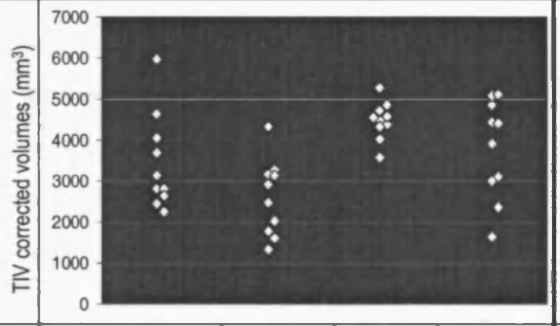
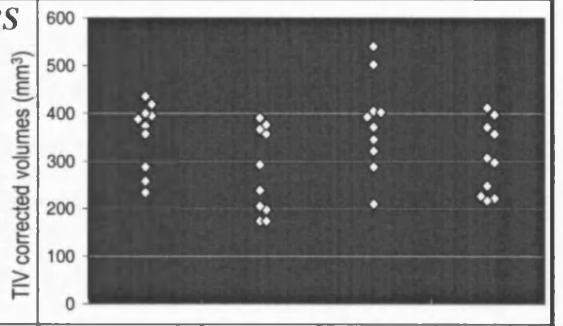
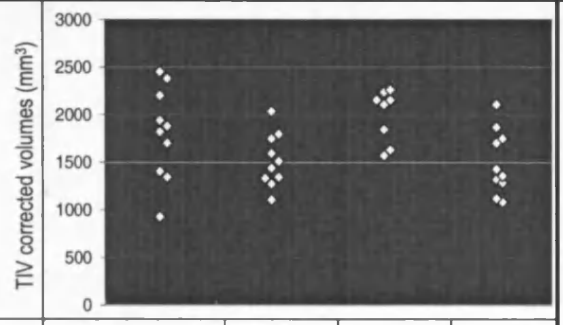
	Left		Right			Left		Right	
	C	AD	C	AD		C	AD	C	AD
Cingulate gyrus 					***PC*** 				
Mean	7687	5613	9361	6845		2202	1260	2463	1344
SD	1738	1591	908	1281		396	621	525	513
Rostral AC 					***RS*** 				
Mean	3292	2459	4459	3589		348	264	366	297
SD	1170	928	458	1210		70	90	97	75
Caudal AC 									
Mean	1741	1479	2006	1468					
SD	481	280	248	337					

Figure 4.3 Volumes of the cingulate gyrus and its subdivisions in controls (C) and AD subjects.
Mean values (SD) and the scatterplots are displayed. PC= posterior cingulate, AC = anterior cingulate, RS = retrosplenial cortex.

There was statistically significant evidence of a group by region interaction (i.e. that the magnitude of the difference between AD-cases and controls differed between regions combining over hemispheres) ($p=0.0025$), a hemisphere by region interaction ($p=0.034$), but no evidence of a group by hemisphere interaction (i.e. no evidence that the case-control difference differed between the left and right side combining over regions) ($p=0.8$). Accordingly, the two-way interaction between group and hemisphere was dropped from the model effectively pooling the differences between AD-cases and controls over left and right sides. All of the cingulate ROIs were significantly smaller in cases compared to controls with the difference in the PC being significantly larger than that in each of the other cingulate regions ($p<0.01$) (see complete group separation on Figure 4.3): rostral AC: 22.5% (95% CI 6.2-35.9%) smaller, $p=0.009$; caudal AC: 20.7% (95% CI 4.1-34.4%) smaller, $p=0.017$; PC: 44.1% (95% CI 32.4-54.8%) smaller, $p<0.001$; RS: 21.5% (95% CI 5.1-35.1%) smaller, $p=0.012$ (Figure 4.3).

On average, cingulate volumes were larger on the right side in all regions in both AD-cases and controls, with some evidence ($p=0.034$) that the magnitude of the left-right difference varies between regions. This left right difference was greatest for the rostral AC and this was the only region where the difference was statistically significant ($p<0.001$).

Despite the fact that mean volumes for all four cingulate ROIs differed significantly between cases and controls, Figure 4.3 shows that there was considerable group overlap which was least in the posterior cingulate ROI. Sensitivity and specificity values for the PC were calculated to determine how well these data might classify individuals into their diagnostic groups. With a cut-off value of $1880mm^3$ for the mean of the left and right PC, sensitivity was 90%, specificity 100%, positive predictive value 100% and the negative predictive value was 91%.

The wide range of MMSEs in the patient group raises the possibility that the results are driven by the very severe cases. To investigate this, Pearson correlation coefficients between MMSE scores and cingulate volumes were calculated. None of the volumes of the cingulate ROIs showed a correlation with MMSE scores that was statistically significant or large enough to support such a hypothesis: rostral AC $r = -0.36$ ($p = 0.38$); caudal AC $r = -0.35$ ($p = 0.39$); PC $r = -0.23$ ($p = 0.59$); RS $r = 0.04$ ($p = 0.93$).

4.2.5. *Discussion*

Using an MR-based manual delineation protocol for the cingulate gyrus all four cingulate ROIs (RAC, CAC, PC, RS) showed significant atrophy in familial AD patients compared with controls. Within the cingulate gyrus the volume reduction was the greatest in the PC ROI. The analysis detected previously described left-right hemisphere asymmetries in cingulate volumes (right>left), which accords with other studies (Paus *et al.*, 1996; Watkins *et al.*, 2001). The severity of the volume loss did not differ between left and right cingulate regions in the AD subjects.

A potential source of variability in region-based manual outlining results from difficulty in accurately correlating cytoarchitectonic borders in histological sections with gross anatomical landmarks on MR-scans *in vivo*. Often the macroscopic sulci do not coincide with borders of microscopically cytoarchitectonical defined cortical areas (Uylings *et al.*, 2005; Zilles, 2004). A high degree of inter-individual variability in sulcal patterns, as is the case with the cingulate gyrus, further adds to this problem (Paus *et al.*, 1996; Vogt *et al.*, 1995). Although manual demarcation of brain regions is time consuming it is still the "gold standard" of region of interest measurement on MRI (Crum *et al.*, 2003; Tisserand *et al.*, 2002). This manual demarcation protocol was standardised as much as possible and yielded high inter- and intra-rater reliability coefficients, ranging from 91.9-99.4%. The landmarks selected to define the borders of the cingulate gyrus and its subdivisions in the present study were based on extensive examination of morphological and cytoarchitectonic studies and the analysed cingulate subregions approximate to Brodmann's areas 24 (rostral + caudal AC), 29+ 30 (RS) and 23 (PC).

Separate evaluation of these regions in AD may be relevant taking into account the fact that neurofibrillary tangles in the RS are present in relatively early neuropathological stages of AD and precede pathological changes in area 23 and 24 (Braak and Braak, 1993). These tangles are not present in healthy controls and therefore may represent pre-clinical stages of AD (Ma *et al.*, 1994). Furthermore, in non-human primates, rostral area 24 and the RS have dense reciprocal connections with the entorhinal and (para)hippocampal cortex, whereas caudal area 24 and area 23 do not (Insausti *et al.*, 1987; Kobayashi and Amaral, 2003; Lavenex *et al.*, 2002; Suzuki and Amaral, 1994). Considering the differences in medial temporal lobe connections and temporal patterns of neuropathological involvement, it is interesting that not the RS but the PC ROI (≈

area 23) showed the greatest volumetric loss in the present study. The apparent selective vulnerability of area 23 in AD cannot be explained as secondary to deafferentiation as a result of medial temporal lobe damage. Rather, this selective volume loss seems to be related to neurodegeneration, since a number of neuropathological studies have demonstrated severe neuron losses in area 23 in those cases with clinically established symptoms of AD (Vogt *et al.*, 1990; Vogt *et al.*, 1998).

The PC ROI was capable of separating AD subjects from controls with a high sensitivity (90%) and specificity (100%). Measurements of the medial temporal lobe have found similar levels of sensitivity and specificity (Juottonen *et al.*, 1999; Killiany *et al.*, 2002; Scheltens *et al.*, 1992). Although the number of subjects in the present study is small, this region seems promising as a diagnostic indicator of AD. Previous macroscopic MRI studies have demonstrated posterior cingulate cortex atrophy in AD patients using semi-automated techniques (Baron *et al.*, 2001; Frisoni *et al.*, 2002a; Scahill *et al.*, 2002). According to these authors, the atrophy appeared to predominate around the splenial sulcus, including the dorsal part of the posterior cingulate gyrus, which partially corresponds with the findings. Atrophy of the rostral (but not the caudal) AC was not evident until later stages of the disease (Frisoni *et al.*, 2002a; Scahill *et al.*, 2002). Automated methods require spatial normalization and smoothing techniques and are probably less suitable for analysing highly variable cortical regions (AC) or small areas (e.g. RS) (Ashburner and Friston, 2000; Crum *et al.*, 2003; Mosconi *et al.*, 2005). A manual region of interest study has shown significant atrophy of the caudal AC and the PC in AD but did not find any significant atrophy in the rostral AC. However, this study included the paracingulate gyrus in the AC and did not subdivide the posterior cingulate gyrus and thus did not assess the retrosplenial cortex separately (Callen *et al.*, 2001). Many of the studies detailed in the literature feature disease groups which have been clinically defined. In this study, there was confirmation of disease in most cases and although the groups are not large, there is a high level of confidence in the diagnosis.

In the present study a manual delineation protocol to measure atrophy of four different cingulate regions on MRI has been described and validated. The present study is the first to find significant atrophy of all four cingulate regions in Alzheimer's disease. The delineation protocol has also been shown to be capable of detecting atrophy in smaller regions of the cingulate gyrus, such as the retrosplenial cortex. To my knowledge, this

study is the first to assess separately the volumes of the posterior cingulate and the retrosplenial cortex. Considering the connectional and functional heterogeneity of these areas, regional cingulate atrophy measures may aid in understanding the way in which the disease begins and progresses and may also help in the search for diagnostic markers of early AD.

4.3. Differences between AD and FTLD in the hippocampus and amygdala

4.3.1. Introduction

In the elderly population AD is the most common cause of dementia (Dawbarn and Allen, 2001). However in patients presenting with dementia under the age of 65, the prevalence of FTLD approximates that of AD (Ratnavalli *et al.*, 2002). Differentiating AD from FTLD is important as prognosis and management is different. A diagnosis of AD is suggested if memory decline is an early clinical feature, while prominent behavioural and language dysfunction suggests a diagnosis of FTLD. However, recent clinicopathological studies have identified cases of FTLD with prominent episodic memory loss (Graham *et al.*, 2005), and a frontal presentation of AD has been described (Galton *et al.*, 2000; Johnson *et al.*, 1999). The overlap in clinical features and diagnostic uncertainty, especially in the early stages has led to interest in imaging to help differentiate FTLD from AD (Koeppe *et al.*, 2005).

Prominent hippocampal atrophy has been demonstrated to be the signature pattern of atrophy on MRI in AD (Callen *et al.*, 2001; Convit *et al.*, 1993; Jack *et al.*, 1997; Krasuski *et al.*, 1998; Lehericy *et al.*, 1994) (see Table 1.1 for a review of cross-sectional measures from the literature), while more anterior temporal losses including severe amygdala atrophy has been suggestive of FTLD (Boccardi *et al.*, 2002; Boccardi *et al.*, 2005; Chan *et al.*, 2001b; Galton *et al.*, 2001). These findings support the clinical presentations since hippocampal damage is linked to loss of episodic memory (Vargha-Khadem *et al.*, 1997) and amygdala damage has been associated with emotional and behavioural dyscontrol (Boccardi *et al.*, 2005; Rosen *et al.*, 2002c).

Pathology remains the "gold standard" in the diagnosis of AD and FTLD. While several MRI studies have examined the ability of volumetric measures to differentiate clinically diagnosed FTLD from AD (Boccardi *et al.*, 2002; Boccardi *et al.*, 2003; Chan *et al.*, 2001b; Frisoni *et al.*, 1999; Galton *et al.*, 2001; Laakso *et al.*, 2000a; van de Pol

et al., 2006) in the absence of histopathological confirmation there is uncertainty in clinical diagnostic accuracy, with a potential for circularity in study design. Therefore, this study assessed the value of hippocampal and amygdala measurements in separating AD and FTLD from controls and from one another, in a completely pathologically defined cohort of AD and FTLD cases.

4.3.2. Subjects and methods: Use of PM/biopsy cases

4.3.2.1.Subjects

Ten pathologically confirmed AD patients, and 17 pathologically confirmed FTLD patients (five Pick's disease (PiD), eight with ubiquitin-only immunoreactive neuronal inclusions (FTLD-U), three tau positive with tau exon 10+16 mutations (tau exon 10^{+16}), and one dementia lacking distinctive histology (DLDH)) were included in this study. Pathological diagnosis was based upon the most recent consensus criteria (McKhann *et al.*, 2001). Pathological confirmation by biopsy rather than post-mortem was available in four of the AD and two of the FTLD subjects.

Medical records were reviewed and the clinical presentation, disease duration and MMSE score (Folstein *et al.*, 1975) at the time of the MRI study were recorded. A family history of dementia was present in all tau exon 10^{+16} cases, two FTLD-U cases and one AD case. The FTLD patients were classified using established clinical criteria into fvFTD (n=12), SD; (n=4) and PNFA (n=1) (Neary *et al.*, 1998). None of the FTLD patients had any clinical evidence of motor neuron disease. Ten healthy age-matched controls without evidence or symptoms of cognitive decline were also included in the study. Subject demographics are reported in Table 4.3.

Table 4.3 Subject Demographics.

	Controls	AD	FTLD
Number	10	10	17
M/F	6/4	7/3	11/6
Age (years):			
mean	55.9	57.0	56.4
(standard deviation)	(11.1)	(8.8)	(10.2)
MMSE (/30):			
mean	28.5	14.7	21.6
(standard deviation)	(1.4)	(5.7)	(6.5)
Sporadic /Familial	NA	9/1	12/5
Mean disease duration (years) at time of scan	NA	3.1*	3.5
		(0.7)	(2.0)
Mean time from scan to death (years) in those coming to post-mortem (6 AD and 15 FTLD)	NA	3.7‡	4.5‡
		(2.3)	(2.1)

Key

* Disease duration not available in one AD patient

‡ data only available for those subjects that came to autopsy (6 AD and 15 FTLD cases)

4.3.3. Image Analysis

4.3.3.1.Image Acquisition

Images were acquired in the standard way as described in Appendix Three. If an individual had more than one diagnostic scan, the earliest scan was chosen for inclusion in the study.

4.3.3.2.Manual Segmentation

The software package MIDAS was used for all image analysis (see page 95). Prior to tracing all scans were registered to a standard template (MNI 305) (Mazziotta *et al.*, 1995). In addition, mirror-image volumetry was used such that traces were performed on the right-hand side of the presented image (see page 101). The operators were blinded to the subject's identity, diagnosis, and left-right orientation of the scan. Amygdala (Whitwell *et al.*, 2005b) and hippocampal (Scahill *et al.*, 2003) measurements were performed (see Appendix 4).

4.3.3.3.Statistical Analysis

Data were analysed using STATA version 8, and SAS (see page 96). Volumes were adjusted for head-size differences using the TIV derived according to a previously described protocol (Whitwell *et al.*, 2001) (see Appendix Four). Standardisation was carried out separately for the amygdala and hippocampus. For both structures this assumed a linear relationship between log transformed values and log TIV with the slopes of the associations estimated from left and right measurements in an expanded control group to improve precision (see Equation 4.1) (Free *et al.*, 1995).

In order to minimise the number of hypothesis tests carried out, a structured analysis was carried out using linear mixed models with interaction terms, and disease group-specific unstructured covariance matrices. Such an approach extends Analysis of Variance by removing the requirement for homogeneity of variances. Such models were used to make comparisons between geometric mean volumes by i) disease group (FTLD vs. AD vs. controls in the primary analysis), ii) amygdala vs. hippocampus and iii) structural laterality. Contrasts of mean levels were used to estimate effect sizes and Wald tests used to test the statistical significance of interaction terms and main effects. Models of this type were also used to make comparisons between geometric mean volumes by FTLD subtype. Because unstructured covariance models do not facilitate comparison of variances between disease groups and regions, direct product covariance matrices (amygdala/hippocampus $\otimes$ left/right side) were used to compare variances between i) disease groups, ii) amydgala and hippocampus and iii) structural laterality. A likelihood ratio test was used to compare the fit of a model where disease group specific amygdala and hippocampal variances were assumed constant with one where they were allowed to be different. A similar approach was used to compare left and right sided variances. In a similar comparison of variances by disease groups left and right side

variances were assumed to be the same. In addition sensitivities for certain specificity cut-offs were calculated to test the relative utility of both structures in discriminating subject groups.

4.3.4. Results

There was no significant difference in age or gender distributions between the FTLD, AD and control groups (Table 4.3). MMSE scores were significantly higher in the control group than both the AD and FTLD groups ($p<0.003$), and significantly higher in the FTLD group than the AD group ($p=0.03$). There was no significant difference in disease duration and time to death between the FTLD and AD groups.

Hippocampal and amygdala volumes from the three subject groups are shown in Table 4.4 and displayed as scatter-plots in Figure 4.4 and Figure 4.5. Relationships between the hippocampus and amygdala in each of the three subject groups on left and right sides are shown in Figure 4.6 and Figure 4.7 respectively. Analysis using linear mixed models on log transformed values showed no evidence ($p > 0.2$ all comparisons) that proportionate differences in volumes between disease groups differed between amygdala and hippocampus and/or between hemispheres, or that proportionate differences in volumes between hemispheres differed between amygdala and hippocampus. Averaging over left and right regions the geometric mean amygdala and hippocampal volumes were respectively 15.0% (95% CI 4.2%, 24.5%) and 16.4% (95% CI 5.9%, 25.6%) lower in AD than controls. In FTLD the equivalent differences were 43.0% (95% CI 31.9%, 52.6%) in the amygdala and 36.1% (95% CI 27.5%, 43.7%) in the hippocampus. The difference in volumes between the FTLD and AD groups was statistically significant ($p<0.01$ in both regions).

There was no evidence ($p>0.2$) that the direct product unstructured covariance matrices did not provide a good fit to the data. Such models demonstrated that variability in hippocampus and amygdala volumes relative to mean levels was greater in the FTLD group than the AD group, and greater in patients than controls ($p<0.001$ for all cases vs. controls; $p=0.02$ for AD vs. FTLD pooling over structures) as suggested by the coefficients of variation reported in Table 4.4. There was no evidence of differences in variability between amygdala and hippocampal volumes (pooling over left and right and disease groups) or between structures on the left and right (pooling over disease groups and structures).

As anticipated from the comparison of means, all four volumes (left and right hippocampus, and left and right amygdala) could significantly discriminate AD (vs. controls), and FTLD (vs. controls and vs. AD) using regression models ($p< 0.05$). When the specificity was set at 80%, the sensitivity for detection of AD subjects from controls ranged from 65% in the right amygdala to 80% in the left amygdala with both hippocampal measures falling within this range. In comparison, the sensitivity for detection of FTLD from controls was over 90% in all structures and reached 100% in the right hippocampus. In addition, for a specificity of 80% the right amygdala could discriminate FTLD and AD subjects with a sensitivity of 88% whilst the sensitivity for the other three structures ranged from 50 to 71%.

Table 4.4 Mean, (95% CI), SD, and CV (100(SD/mean)) of volumes of TIV adjusted hippocampus and amygdala in controls, AD and FTLD subjects.*

Volume mm^3	**Controls**		**AD**		**FTLD**	
	Left	**Right**	**Left**	**Right**	**Left**	**Right**
<u>Amygdala</u>						
Mean (95% CI)	1555 (1491, 1619)	1618 (1514, 1721)	1314 (1128, 1500)	1415 (1258, 1573)	940 (747, 1133)	976 (830, 1122)
SD	89	144	260	221	375	283
CV	5.7	8.9	19.8	15.6	39.9	29.0
<u>Hippocampus</u>						
Mean (95% CI)	2898 (2723, 3073)	2967 (2727, 3207)	2410 (2088, 2732)	2534 (2276, 2792)	1883 (1643, 2123)	1962 (1722, 2203)
SD	245	336	449	361	466	468
CV	8.5	11.3	18.6	14.2	24.7	23.9

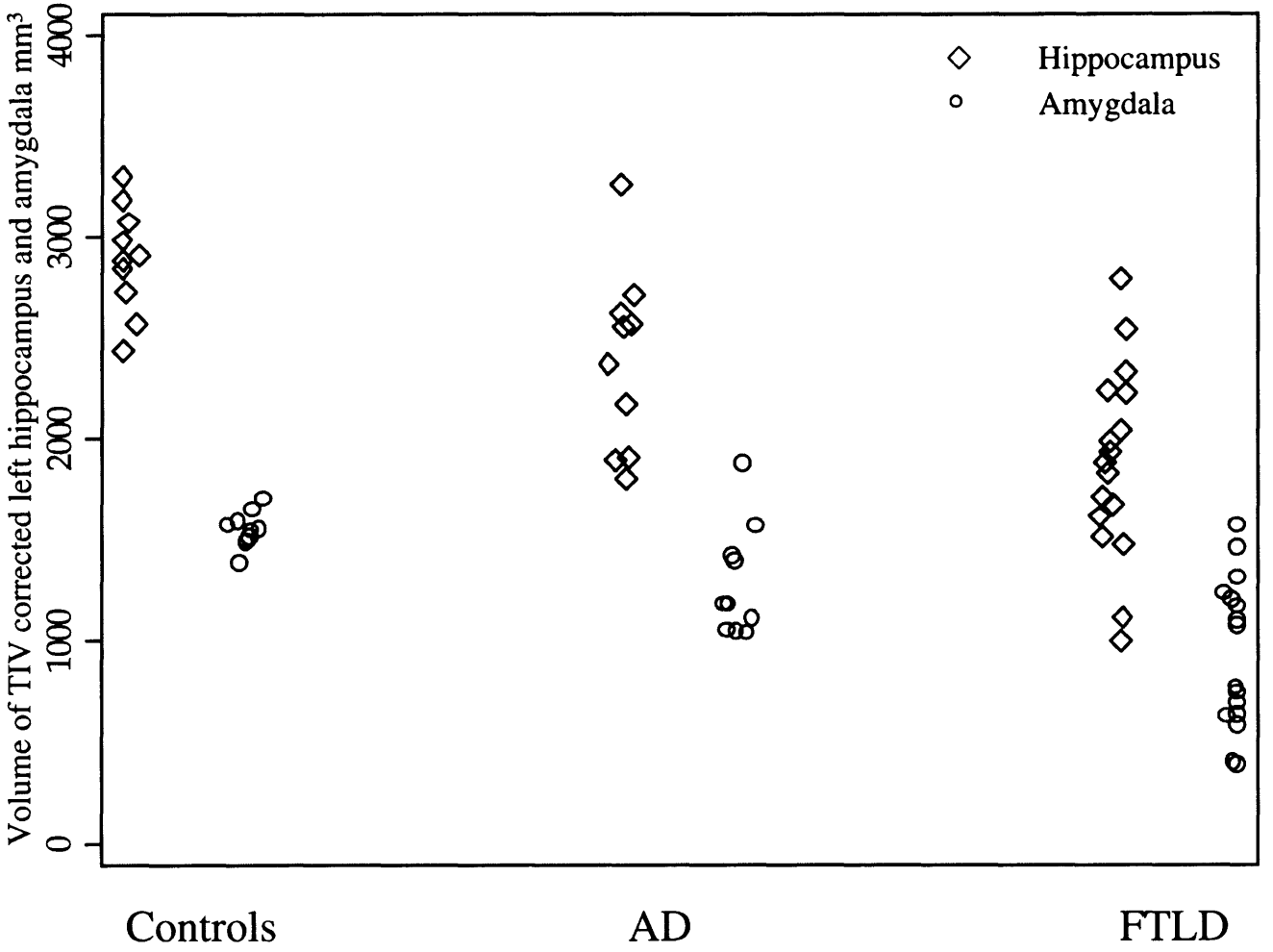


Figure 4.4 Left amygdala and hippocampal volume in controls AD and FTLD.

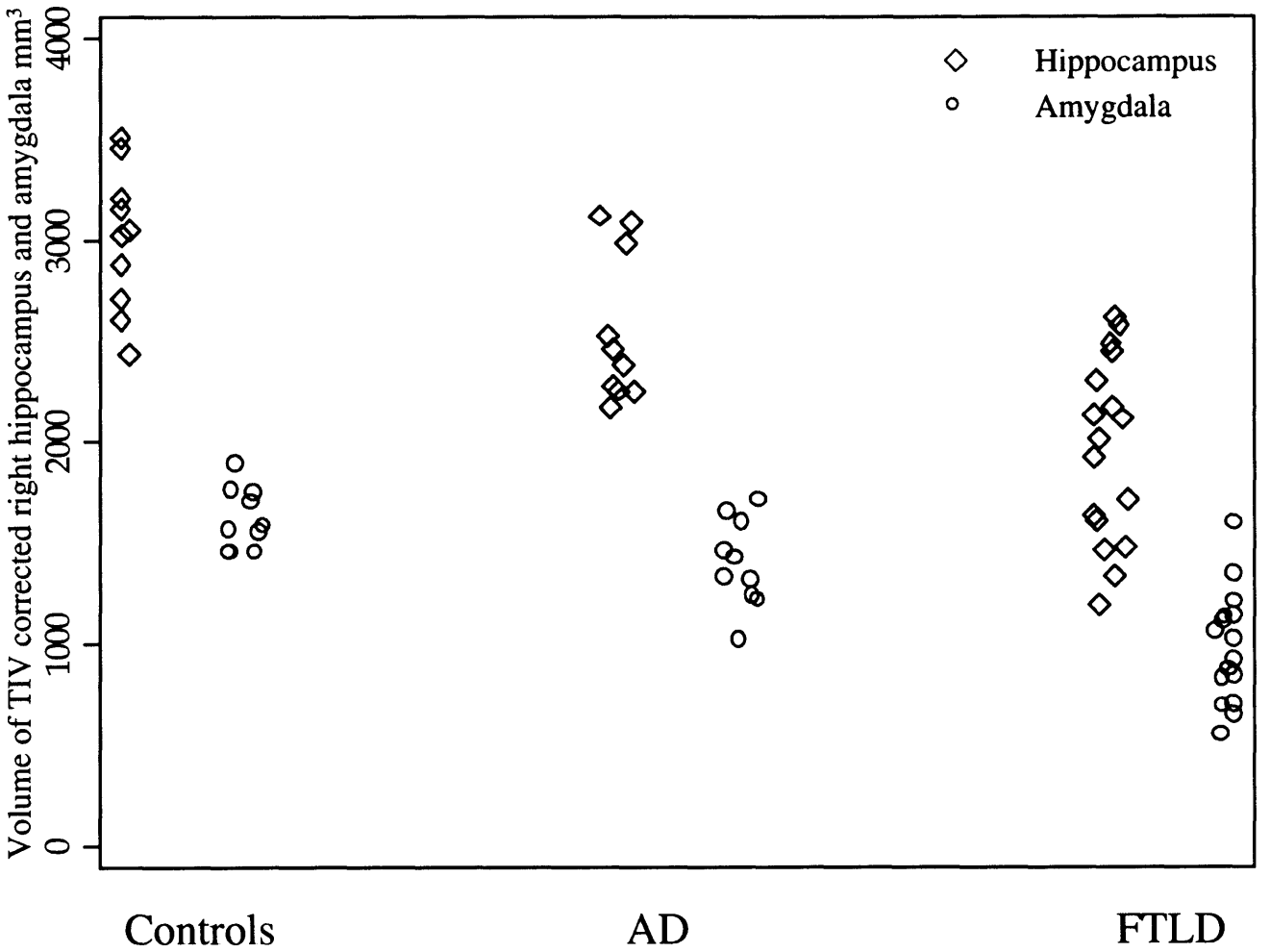


Figure 4.5 Right amygdala and hippocampal volume in controls AD and FTLD.

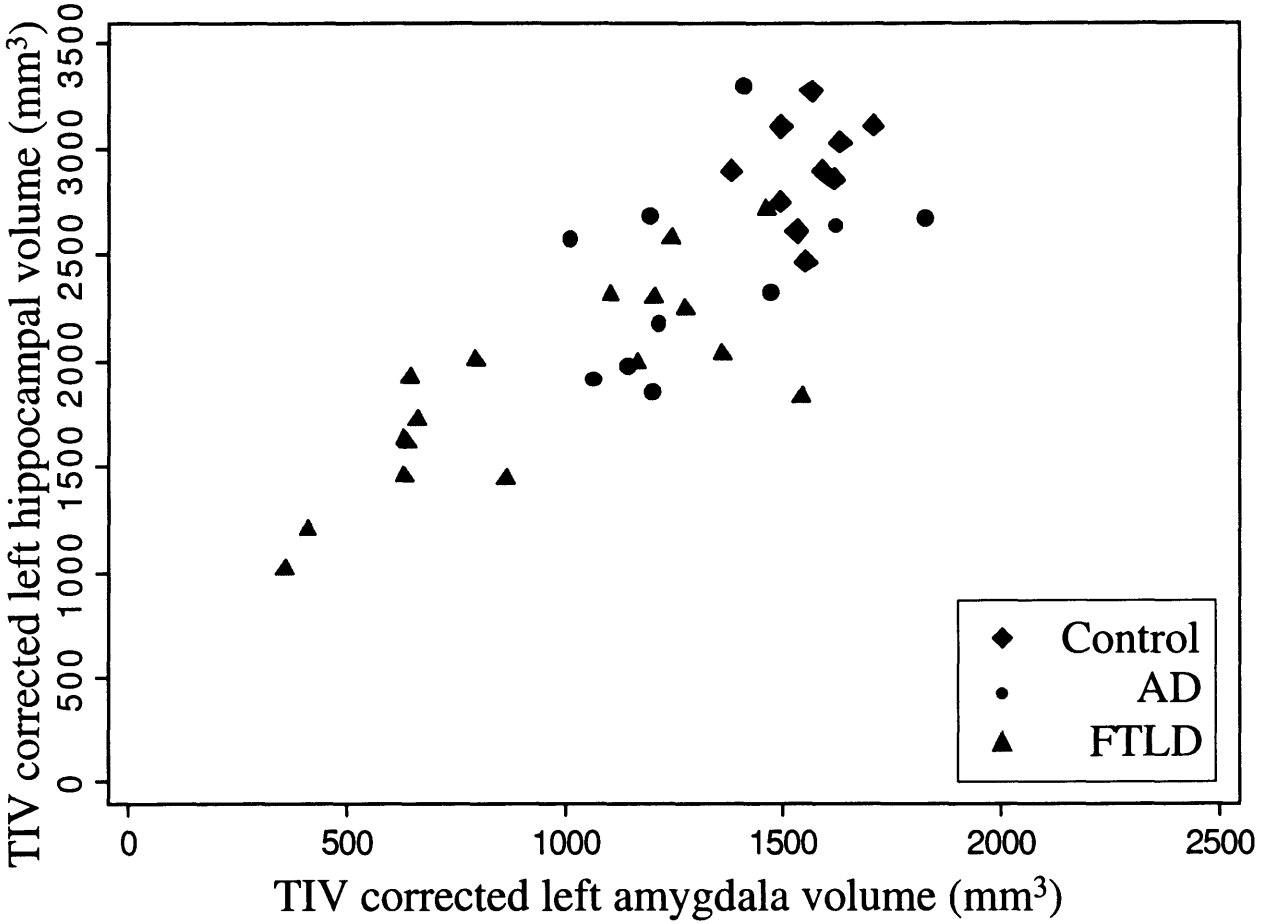


Figure 4.6 Left hippocampal and amygdala volume in controls, AD and FTLD.

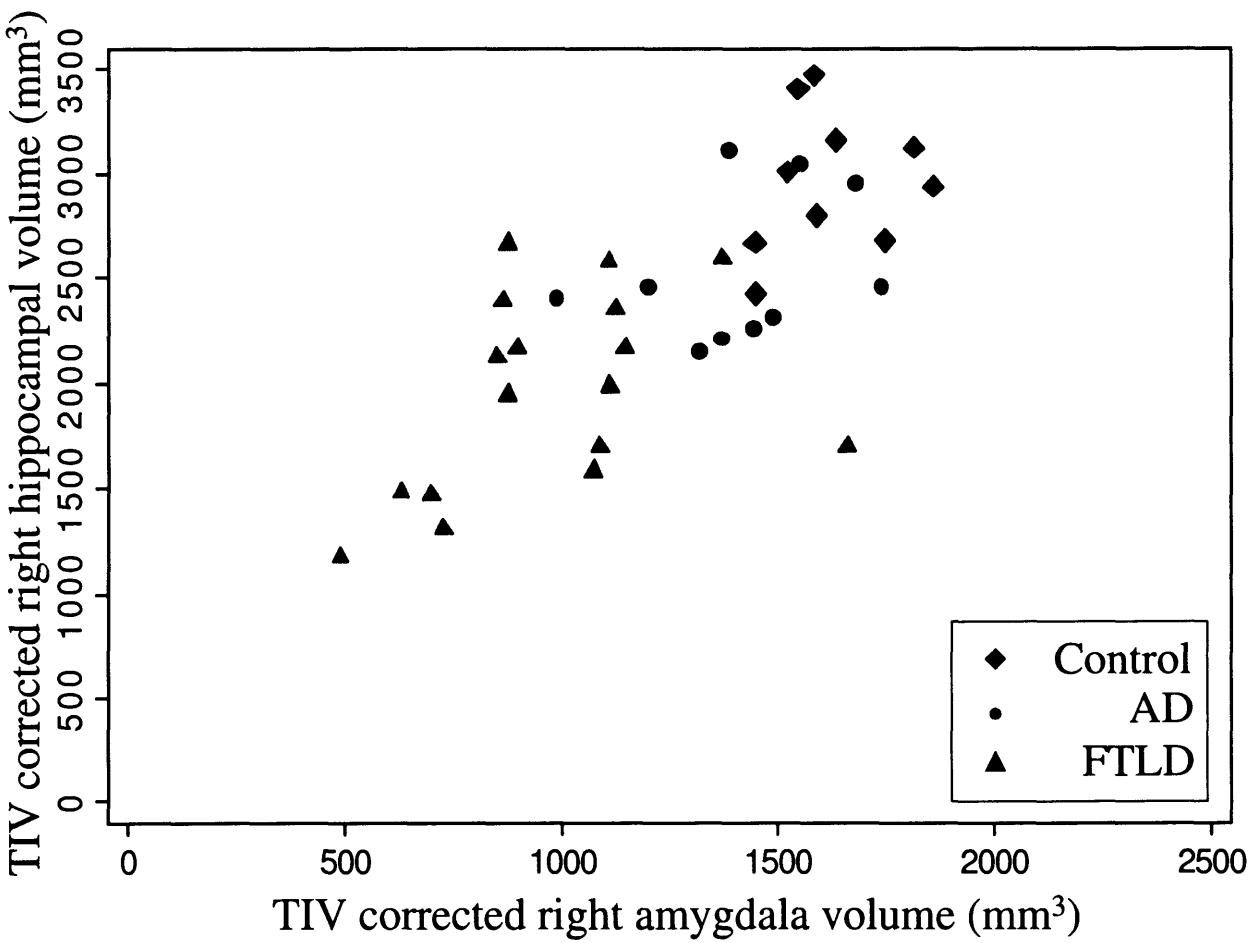


Figure 4.7 Right hippocampal and amygdala volume in controls, AD and FTLD.

4.3.4.1.Subgroups of FTLD

The FTLD group was subdivided pathologically into FTLD-U, tau exon 10^{+16}, PiD and DLDH. Owing to the presence of only one subject in the DLDH group, this group was removed from the pathological subgroup analysis. There was no evidence for differences in mean hippocampal or amygdala volumes between the three remaining groups ($p>0.5$).

The FTLD group was also subdivided clinically into fvFTD, SD and PNFA. The amygdala and hippocampal volumes in these subgroups are shown in Figure 4.8 and Figure 4.9. Owing to there being only one individual in the PNFA group, this group was removed from further analysis. There was evidence that the ratio of geometric mean volumes in the two subtypes differed between the amygdala and hippocampus ($p=0.018$) and by hemisphere ($p=0.0004$). On average the left hippocampus was 14% smaller in SD than in fvFTD, whilst the right hippocampus was 37% larger. On average the left amygdala was 39% smaller in SD than in fvFTD, whilst the right amygdala was only 1% smaller. These differences can be seen graphically in Figure 4.10 and Figure 4.11 which represent area per slice along the length of the union of the hippocampal and amygdala regions. These profiles also show the relatively greater involvement of anterior regions compared with posterior regions in both fvFTD and SD, compared to AD and controls.

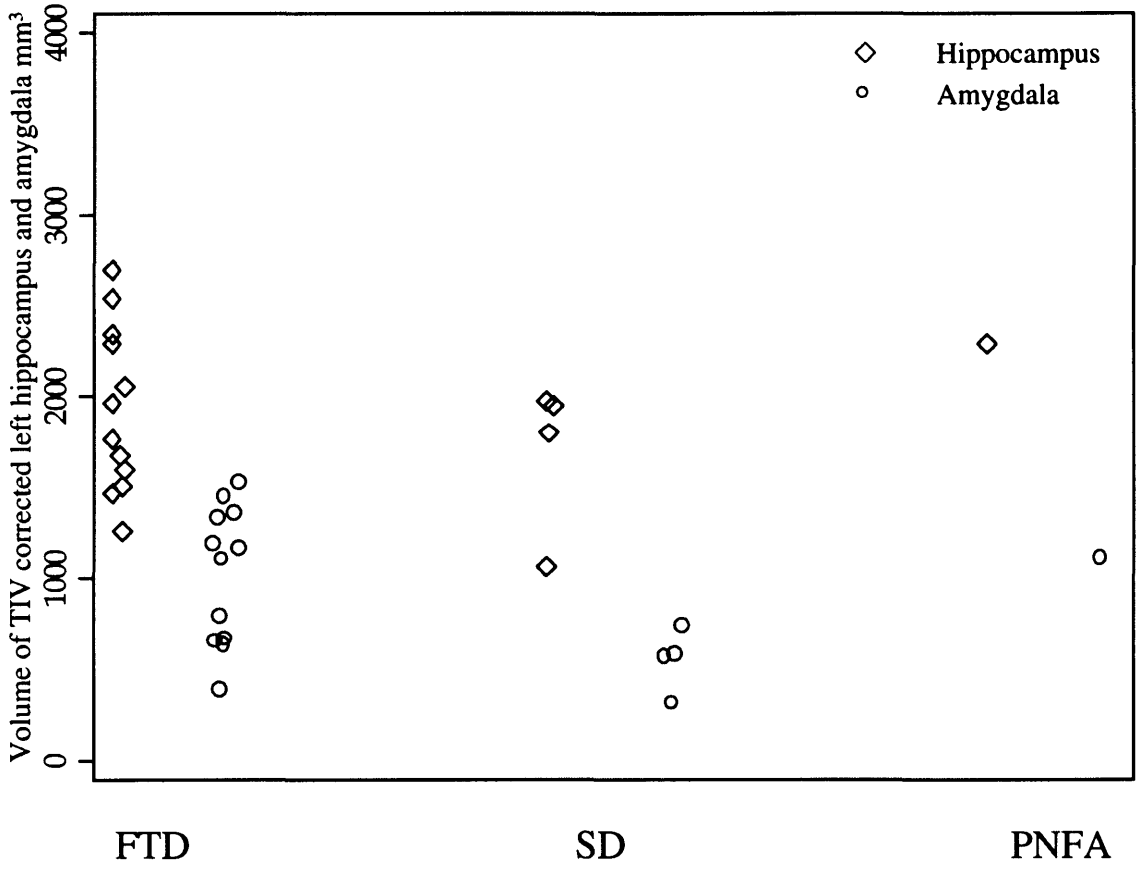


Figure 4.8 Volumes of left hippocampus and amygdala in clinical subtypes of FTLD.
FTD is the frontal variant of FTLD, SD is semantic dementia and PNFA is progressive non-fluent aphasia.

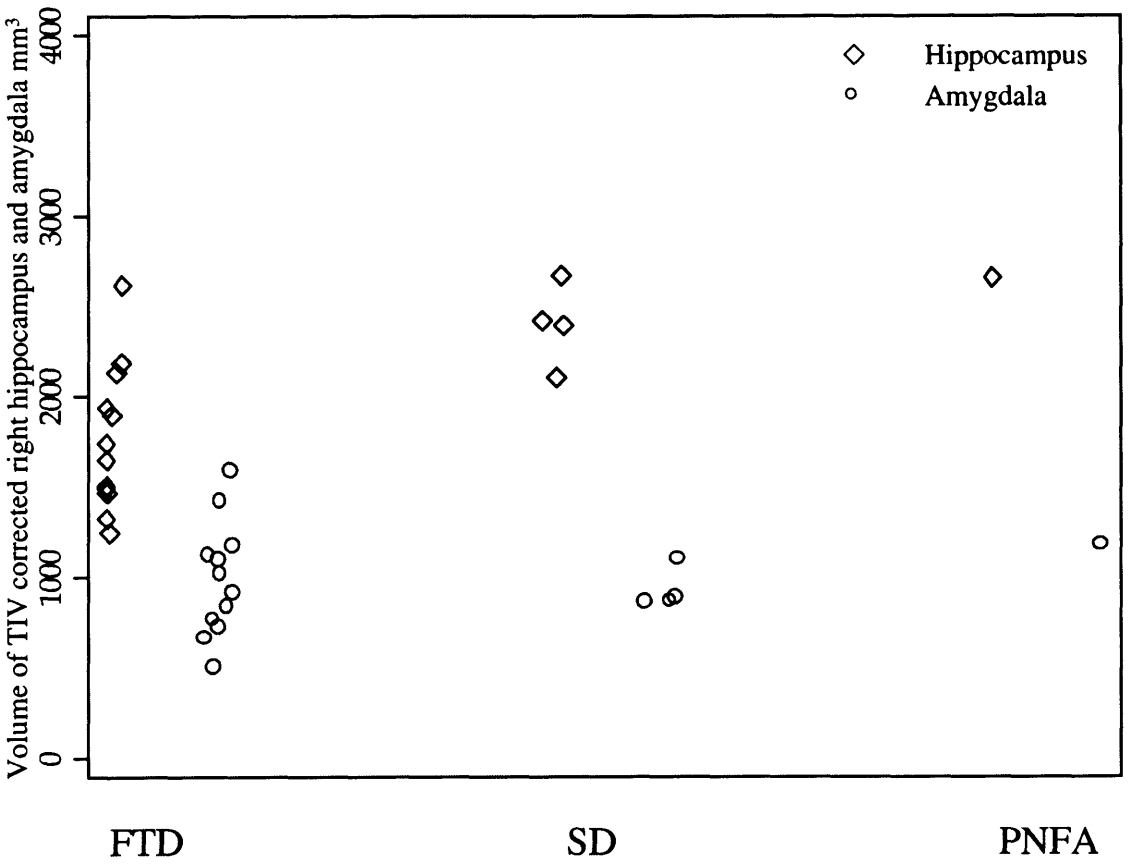


Figure 4.9 Volumes of right hippocampus and amygdala in clinical subtypes of FTLD.
FTD is the frontal variant of FTLD, SD is semantic dementia and PNFA is progressive non-fluent aphasia.

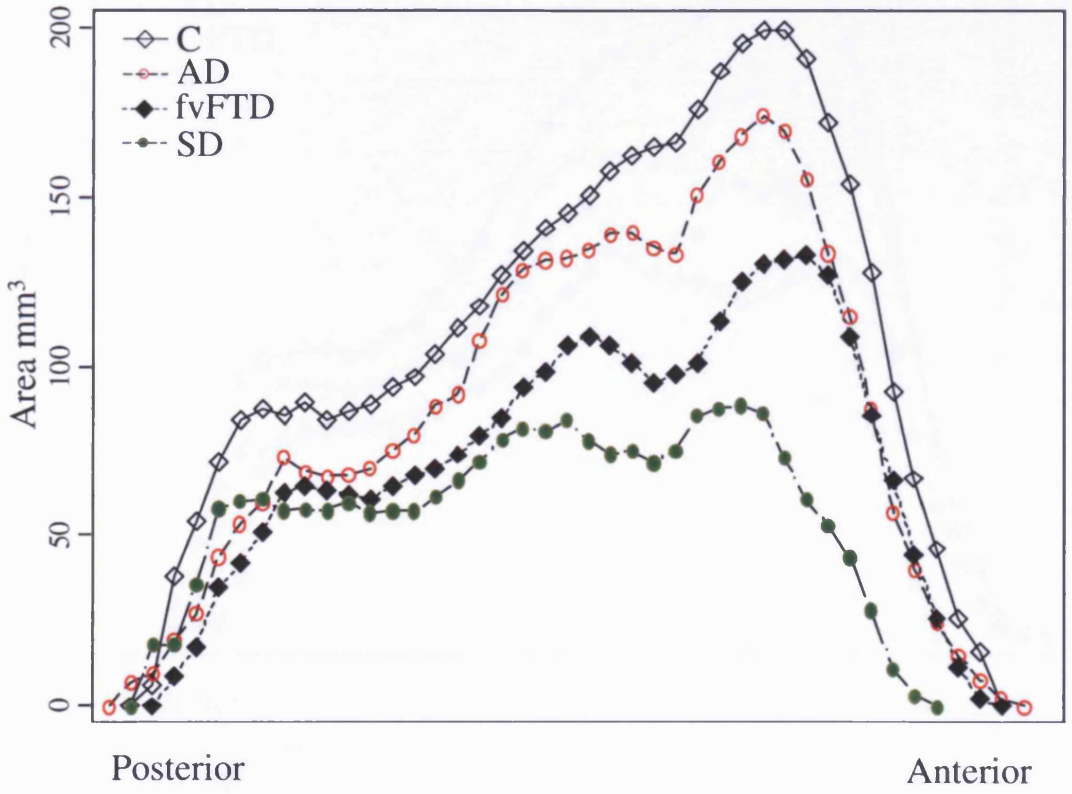


Figure 4.10 Left amygdalo-hippocampal complex area profiles along the posterior to anterior axis (sagittal view) in controls, AD and two clinical phenotypes of FTLD.

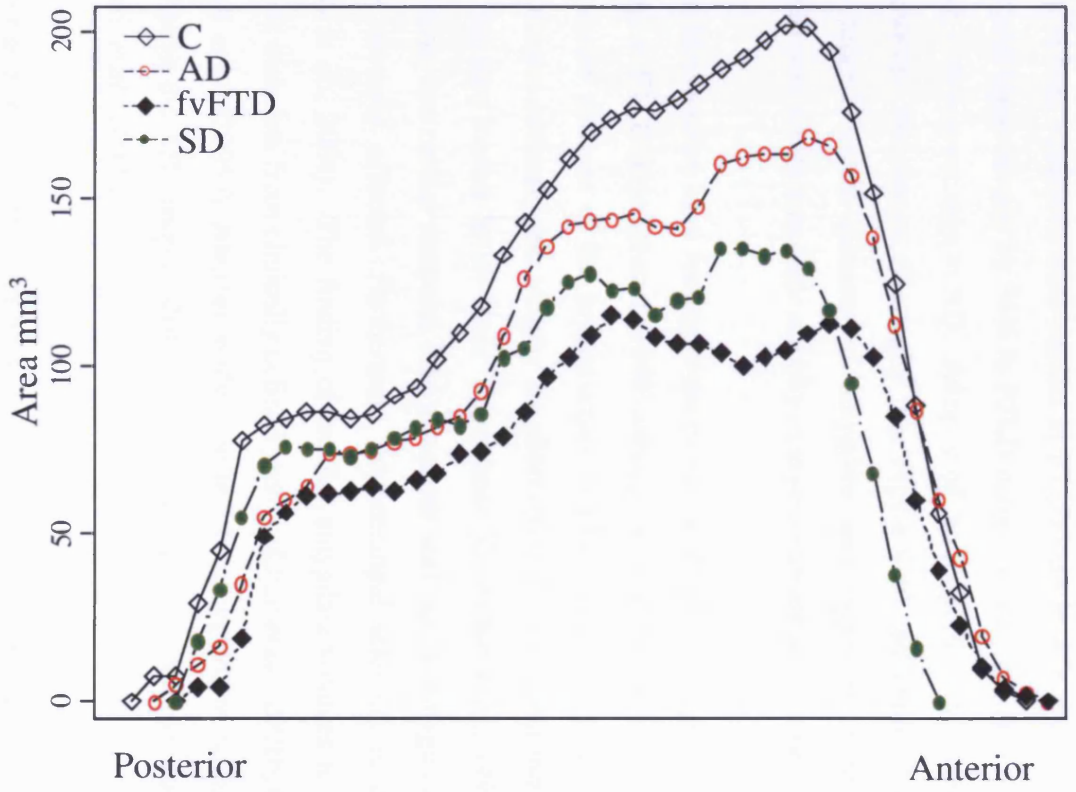


Figure 4.11 Right amygdalo-hippocampal complex area profiles along the posterior to anterior axis (sagittal view) in controls, AD and two clinical phenotypes of FTLD.

4.3.5. Discussion

This study examined the patterns of amygdala and hippocampal atrophy in pathologically confirmed cases of FTLD and AD. Volumes of both structures were reduced in both the FTLD and AD groups compared with controls, consistent with previous studies (Boccardi *et al.*, 2002; Boccardi *et al.*, 2003; Callen *et al.*, 2001; Chan *et al.*, 2001b; Frisoni *et al.*, 1999; Jack *et al.*, 1997; Krasuski *et al.*, 1998; Lehericy *et al.*, 1994). In addition, the volumes of both structures were smaller in FTLD than in AD; the amygdala was reduced by 43% and hippocampus by 36% in FTLD compared with controls, compared with only 15% and 16% respectively in AD. Atrophy of the hippocampus is clearly not specific to AD. However, the pattern of atrophy (see Figure 4.10 and Figure 4.11) varies between groups. Indeed, severe patterns of amygdala and hippocampal atrophy especially if associated with severe amygdala atrophy or asymmetry are suggestive of FTLD.

Previous MRI studies have found greater or similar degrees of hippocampal atrophy in AD compared to FTLD (Frisoni *et al.*, 1999) although atrophy was predominantly located in the more anterior portions of the hippocampus in FTLD (Laakso *et al.*, 2000a). However, severe patterns of focal lobar atrophy are observed in FTLD at post-mortem; gyri are often so thin that they have a "knife-blade" appearance (Gustafson *et al.*, 1992). So it is perhaps unsurprising that medial temporal lobe structures such as the amygdala and hippocampus may be severely affected. Furthermore, hippocampal sclerosis is common in FTLD (Josephs *et al.*, 2004). The finding of smaller amygdala volumes in FTLD than AD is consistent with data from clinically defined groups (Chan *et al.*, 2001b; Galton *et al.*, 2001; Whitwell *et al.*, 2005b). Another study reported the amygdala to be larger in FTLD than AD although the AD group in that study was nearly a decade older than the FTLD group (Boccardi *et al.*, 2002).

Measurements of both the hippocampus and amygdala provided good discrimination of FTLD and AD subjects from controls. The specificity of 80% and sensitivity of 75% for discrimination of AD from controls using the hippocampal (left) volume is similar to previous hippocampal studies (Xu *et al.*, 2000). The discrimination of FTLD subjects from controls was achieved with a specificity of 80% and sensitivities greater than 90% for all structures, reaching 100% for the right hippocampus, which is superior to that found in

some studies (Frisoni *et al.*, 1999). Comparisons between studies are difficult however because of differences between subject groups (especially since pathology is often lacking) and differences in severity of disease at the time of the scan. Previous studies have found poor discrimination between AD and FTLD using single volumetric measures (Frisoni *et al.*, 1999; Whitwell *et al.*, 2005b). This study however suggested that for a specificity of 80% both the right hippocampus and amygdala could differentiate FTLD from AD with sensitivities over 70%. This improved discrimination may be a result of this study utilising a pathologically confirmed cohort, as recent studies suggest that only between 63 and 79% of pathologically-confirmed FTLD cases are correctly classified clinically on their first assessment (Knopman *et al.*, 2005; Rosen *et al.*, 2002b).

In the FTLD subgroup analysis, a clear differential pattern of amygdala and hippocampal atrophy by pathological subtype was not found. This is consistent with a previous study that showed medial temporal lobe atrophy in all three pathological subtypes (Whitwell *et al.*, 2005a). However, when dividing the FTLD group by clinical subtype, marked differences in the patterns of atrophy were found. First, the SD patients showed a more asymmetrical pattern of atrophy than the fvFTD group with much greater atrophy on the left, and second, the SD group had disproportionately smaller amygdala volumes compared with the hippocampus than the fvFTD group. This is consistent with previous volumetric studies on clinically diagnosed patients that have shown asymmetric temporal lobe atrophy, with particularly severe amygdala atrophy, in SD patients (Chan *et al.*, 2001b; Galton *et al.*, 2001; van de Pol *et al.*, 2006; Whitwell *et al.*, 2005b), yet relatively symmetrical atrophy in patients with fvFTD (Boccardi *et al.*, 2002; Rosen *et al.*, 2002a; van de Pol *et al.*, 2006). These findings support the idea that the clinical phenotype in FTLD reflects the pattern of tissue loss more closely that the molecular pathology of the diseases. Both SD and fvFTD also showed relatively greater involvement of anterior than posterior regions of the amygdalo-hippocampal structure. This anterior-posterior gradient has been previously reported in SD patients (Chan *et al.*, 2001b).

The main strength of this study is the pathological confirmation of disease in all cases. However, a limitation of using a pathologically-defined patient group is that the study may be biased towards cases which were difficult to diagnose during life. Another limitation

with all studies of this type is that of accurately matching AD and FTLD patients for disease severity. Although the MMSE has been useful as a measure of disease severity in AD, it is a less comparable measure in FTLD. Owing to this potential problem the disease duration of all patients was also assessed. Both the AD and FTLD groups had similar mean disease duration of approximately three years. Finally, the patient groups were also relatively small in this study. Therefore more subtle differences that may exist in hippocampal and amygdala atrophy between AD and FTLD, and between the pathological and clinical subgroups of FTLD, may not have been detected.

In conclusion, this study has shown that both the hippocampus and amygdala are reduced in volume in pathologically confirmed cases of AD and FTLD compared with controls, with the greatest loss shown in the FTLD subjects. This demonstrates firstly that hippocampal atrophy is not specific to AD, and secondly that severe or asymmetrical patterns of amygdala and hippocampal atrophy suggest pathological FTLD. Both hippocampus and amygdala provide good discrimination of disease groups from controls, and between disease groups.

4.4. Chapter conclusions

This chapter describes the development of new segmentation protocols and their application to a small group of Alzheimer's disease and control subjects at one imaging time-point. The chapter further describes the application of existing protocols in a group of post-mortem confirmed AD, FTLD and age-matched controls. Although there are limitations with this type of single time-point analysis, principally as it is confounded by inter-individual variability, this chapter demonstrates that the use of cross-sectional measures is reliable and can find differences in volumes of structures in different diseases. For diagnosis, the ideal situation is that subjects only have scanning at one time-point as diagnosis can potentially be reached quickly, and therefore it is important to understand the utility and limitations of such imaging techniques.

5. LONGITUDINAL REGIONAL CHANGE USING MANUAL SEGMENTATION

5.1. Chapter introduction

Adding another dimension to cross-sectional change

As assessed in Chapter 4, cross-sectional volumetric measures of regions of interest can reveal patterns of atrophy which may i) distinguish between different diseases (Chan *et al.*, 2001b) ii) give quantitative information regarding the areas of the brain that have degenerated prior to the acquisition of the MR scan and iii) predict which subjects may clinically progress in the future (Jack *et al.*, 2000; Killiany *et al.*, 2000). However there are a number of problems associated with cross-sectional analysis. First, there is a large degree of inter-individual variation which confounds the interpretation of individual results even following the application of procedures which allow adjustments for head size (Jack *et al.*, 1998). Secondly, information from one time-point gives no real or accurate information regarding the progression of disease and the rate of this change. Four dimensional analyses (the additional dimension being time) of regions of interest can give this information and therefore aid in the understanding of the underlying disease processes and provide potential markers of disease progression.

The "gold-standard" technique for measuring longitudinal change

Much like cross-sectional analysis, the "gold standard" of longitudinal analysis of substructures of the brain is manual delineation; whereas whole brain change is often calculated using semi-automated techniques (Fox and Freeborough, 1997; Jack *et al.*, 2004; O'Brien *et al.*, 2001). The majority of published results of longitudinal brain substructure change have applied cross-sectional delineation protocols to both baseline and follow-up scans (Du *et al.*, 2004; Jack *et al.*, 1998; Jack *et al.*, 2004; Scahill *et al.*, 2003) (see Table 1.3-Table 1.5 for a list of studies). The change in volume is then calculated as an annualised proportional loss compared with the original baseline volume. Manual delineation is considered to be the best technique of detecting change in substructures as many of these, such as the hippocampus, are small in volume, complex in shape and difficult to delineate automatically; the hippocampus has borders which are a mixture of different tissue types (see Figure 8.10). Manually-generated longitudinal results also give a comparison with which new semi-automated or automated techniques can be compared.

Regions of interest

As outlined in Chapter 1 (page 37, and Table 1.3-Table 1.5), many of the areas which have been studied longitudinally are those which have initially been studied cross-sectionally. These regions were originally targeted by researchers as they were found to be affected by the disease determined by: visualization of structural imaging, reduced metabolism or function measured by functional imaging (Valla *et al.*, 2001) or altered metabolite levels detected by MRS (Kizu *et al.*, 2004), or have been found to be affected pathologically (Braak *et al.*, 1993). Changes in the volume of brain regions studied on serial volumetric MR can either be measures of global change such as whole brain change (Table 1.2) or ventricular change (Table 1.4), or measure change in specific substructures such as the hippocampus and entorhinal cortex (see Table 1.3 and Table 1.5 respectively).

Limitations of these techniques

These measures have potential use and as such have been used in trials as surrogate markers of progression (Fox *et al.*, 2005; Jack *et al.*, 2003; Krishnan *et al.*, 2003) in addition to many large studies (see Table 1.3). However, manual outlining is labour-intensive (often months of prior training is required), subject to human error and requires highly skilled operators; often the level of change calculated within the region approximates to the level of error genenerated from uncorrelated errors in the delineation of matched baseline and repeat regions. As a result the use of these techniques in a clinical situation is limited to specialist referral centres and exceptional cases.

Chapter outline

This chapter describes the application of manual delineation protocols to specific disease groups in order to answer questions regarding i) differential four dimensional patterns of atrophy in sub-regions of the brain known to be affected by specific diseases, and ii) the intervals over which change can be measured in the hippocampus in AD.

5.2. *Application of cingulate ROI protocol to longitudinal studies*

5.2.1. *Introduction*

As discussed in Chapter 4 (see page 97) cross-sectional volumetric MRI studies of the cingulate gyrus have revealed significantly lower cingulate volumes in AD compared with controls (Callen *et al.*, 2001; Killiany *et al.*, 2000). This concurs with voxel-based morphometric (VBM) data which also reveals reduced grey matter density in the cingulate gyrus in AD, more specifically in the posterior region (Baron *et al.*, 2001; Frisoni *et al.*, 2002a; Scahill *et al.*, 2002). Magnetic resonance spectroscopy (MRS) in AD patients shows metabolite changes in the posterior section of the cingulate gyrus compared with controls (Kizu *et al.*, 2004). PET and SPECT studies have also revealed hypometabolism in the cingulate gyrus in AD compared with controls, especially in the posterior section (Matsuda *et al.*, 2002; Minoshima *et al.*, 1997; Valla *et al.*, 2001). In FTLD, there have been fewer studies; VBM has shown predominant decrease in grey matter volume of the posterior cingulate (Whitwell *et al.*, 2004a) and less cingulate atrophy in the posterior section of the cingulate with relatively more anterior temporal change compared with AD (Boxer *et al.*, 2003), and MRS has shown metabolite changes in the posterior cingulate region compared with controls (Kizu *et al.*, 2004).

The hippocampus has been more extensively studied as discussed in the previous chapters (see page 113 and Table 1.1 and Table 1.3). Structural imaging based studies (mainly MRI) have shown increased atrophy of the hippocampus in established AD (Chan *et al.*, 2001b; de Leon *et al.*, 1994; Jack *et al.*, 1998), and in individuals with presymptomatic familial AD (Fox *et al.*, 1996b; Schott *et al.*, 2003) or subjects with mild cognitive impairment (MCI) at higher risk of progressing to AD (Jack *et al.*, 2004). Autopsy studies in AD suggest that hippocampal atrophy, independent of neurofibrillary tangles, predicts memory deficits (Mortimer *et al.*, 2004). As discussed and shown in the previous chapter (see page 113), the hippocampus may also be severely affected by the progression of FTLD with increased atrophy of this structure particularly seen in FTLD variants with temporal lobe predominance (e.g. semantic dementia) (Frisoni *et al.*, 1999; Laakso *et al.*, 2000a). Hippocampal atrophy in semantic dementia may be at least as severe as in AD although usually with a more anterior focus (Chan *et al.*, 2002).

As discussed in the previous chapter (see page 113) at early stages of AD and FTLD it can be difficult to differentiate between the two diseases clinically and although this may be easier as the diseases progress, both diseases may be misdiagnosed. Assessment of MRI scans might aid this differential diagnosis and provide information about the progression of these diseases. In order to establish whether volumetric measures are truly useful in distinguishing disease groups from each other, studies should be ideally performed using scans of subjects with pathologically-confirmed disease as clinical diagnosis is not 100% accurate.

This study aimed to establish the rates of atrophy of the cingulate and hippocampus in confirmed AD and FTLD compared with controls. The secondary aim was to establish the distribution of atrophy rates along the length of the cingulate gyrus. The hypothesis was that these may be different between groups with more anterior loss in FTLD.

5.2.2. *Subjects and methods*

5.2.2.1.Subjects

Nineteen AD patients (ten pathologically, eight genetically confirmed by chromosomal analysis and one autosomal dominant inheritance of AD with post-mortem in a first degree relative), eight FTLD patients (six pathologically confirmed and two from autosomal dominant pedigrees with pathology in a first degree relative), and 11 healthy controls of a similar age (one of which had normal brain reported at post-mortem) were included in this study. Subject demographics are reported in Table 5.1. Where both genetic and pathological confirmation of disease was obtained, patients were included in the pathological group. None of the patients or controls were on medication that would be likely to influence these results. One control subject was normotensive on treatment for hypertension. None of the controls had any diseases that may have influenced these results. An exclusion criteria for the controls was the presence of neurological or psychiatric morbidity. Hypertension was not an exclusion factor.

Within the FTLD group a mixture of pathologies were reported at post-mortem. Three patients had FTLD-U: ubiquitin positive, tau and alpha-synuclein negative, abnormal neuritis and neuronal, intra-cytoplasmic inclusions in the frontal cortices or the hippocampal dentate fascia. Three patients had a positive family history of FTLD with tau exon 10^{+16} C-T (cytosine to thymine) splice site mutations. Of those remaining with post-mortem in a family member one had tau exon 10^{+16} and one had FTLD-U. With respect to clinical presentation, five of the group were of the frontal variant, two had semantic dementia and one had progressive non-fluent aphasia.

5.2.2.2.Image Acquisition

All subjects had two volumetric T1-weighted studies acquired one-two years apart in a standard manner (see Appendix Three).

Table 5.1 Subject Demographics.

	Controls	AD	FTLD
Number	11	19	8
M/F	5/6	10/9	3/5
Age (years):			
Mean	56.0	56.3	55.9
(SD)	(14.3)	(10.6)	(7.6)
MMSE (/30) at first scan:			
Mean	29	20	24
(SD)	(1)	(7)	(6)
Sporadic /Familial	NA	8/11	2/6
Pathological /Genetic confirmation of status	1/0	11/8	6/2
Scan interval (days):			
Mean	430	567	453
(SD)	(238)	(292)	(181)
Interval between mid-point of scans and death (years):			
Mean	NA	3.9*	4.9¥
(SD)		(1.9)	(0.7)

(Data available in 8 cases; ¥ data available in 6 cases)*

5.2.2.3.Manual Segmentation

Scans were transferred to a Sun workstation. The software package MIDAS was used for all manual segmentation (Freeborough *et al.*, 1997). The boundaries of the structures were traced with two orthogonal views visible.

The cingulate was defined as the grey matter of the cingulate gyrus including the retrosplenial cortex (approximating to Brodmann areas 23,24,24',29 and 30) (Vogt *et al.*, 2001). Measurements were taken from sagittal slices from the medial to lateral boundaries and this region was edited in the coronal view. The posterior limit of the cingulate gyrus was taken as the splenial sulcus, the superior and anterior limit as the bottom of the cingulate sulcus and the inferior limit as the corpus collosum. Using the brain mask in MIDAS, a consistent threshold of 70% and 110% of mean brain intensity was applied to the ROI to exclude lower intensity voxels which were predominantly CSF, and higher intensity voxels that were predominantly white matter. See Figure 5.1, and for a full anatomical description, see Chapter 4, page 103.

The hippocampus was defined as including the hippocampus proper, the subiculum and the alveus. Measurements were taken from coronal slices from the posterior to anterior boundaries using a standard neuroanatomical atlas (Duvernoy, 1998). The posterior limit of the HF was defined as the coronal slice where the longest length of the crus of the fornix was seen (Watson *et al.*, 1992). The hippocampus was bounded superiorly, medially and laterally by CSF and inferiorly by the white matter of the subiculum. The head of the hippocampus was delineated from the amygdala by inclusion of the alveus, which was best seen as a band of high signal intensity on the sagittal sections. A consistent threshold of 70% of mean brain intensity was applied to exclude voxels with a lower intensity which were predominantly CSF. Again this mean brain intensity was calculated automatically using the MIDAS generated brain mask, see Figure 5.2 and Appendix Four.

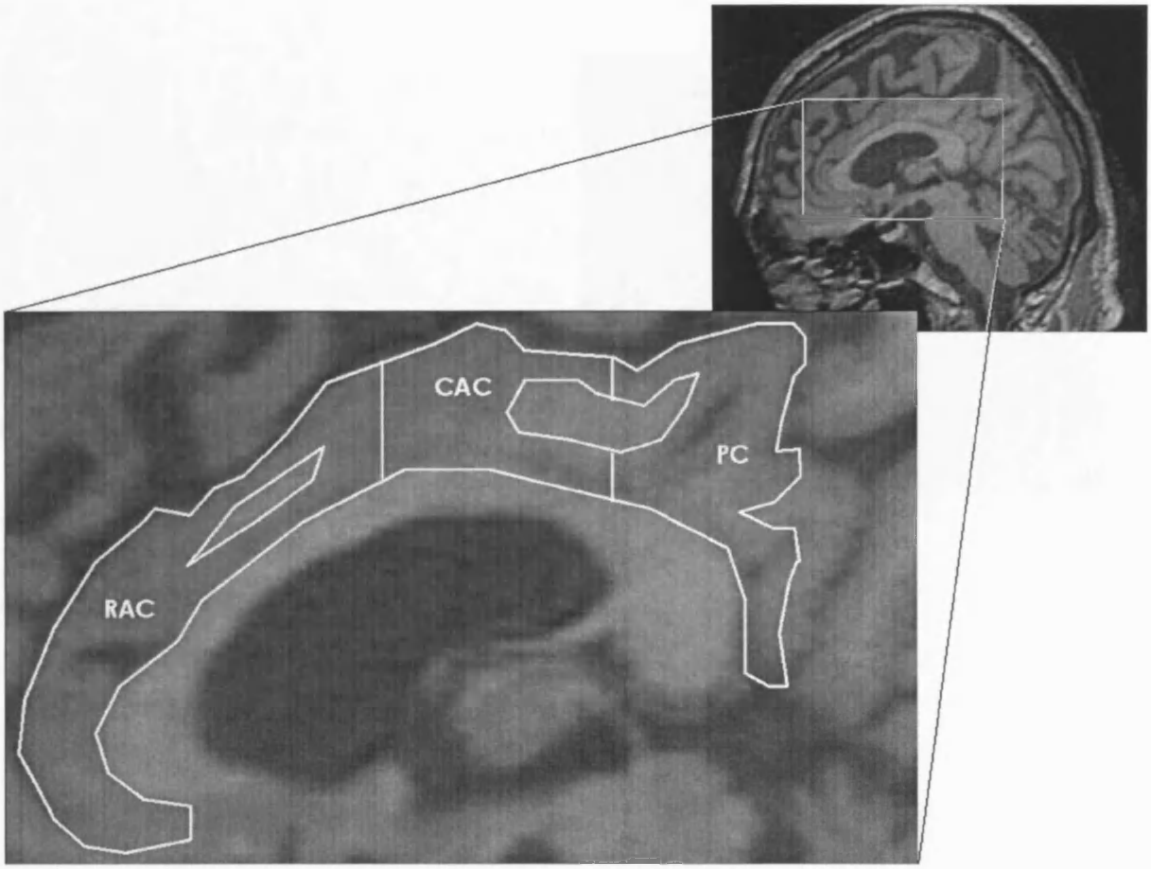


Figure 5.1 A schematic representation of the cingulate gyrus and its subdivisions. The cingulate is outlined in the sagittal view, from medial to lateral parts using strict relative-to-brain thresholds of 70% and 110% to aid delineation of grey matter from CSF and white matter respectively. This region is then edited coronally. Small numbers of partial volume voxels (which may contain both CSF and grey matter) may be included in the segmentation by use of these thresholds. Subdivisions are manually determined using the anterior and posterior comissures as cut-offs. Cingulate gyrus anterior to the anterior commissure is considered to be rostral anterior cingulate (RAC), cingulate posterior to the posterior commissure is posterior cingulate (PC) and cingulate between these regions is caudal anterior cingulate (CAC).

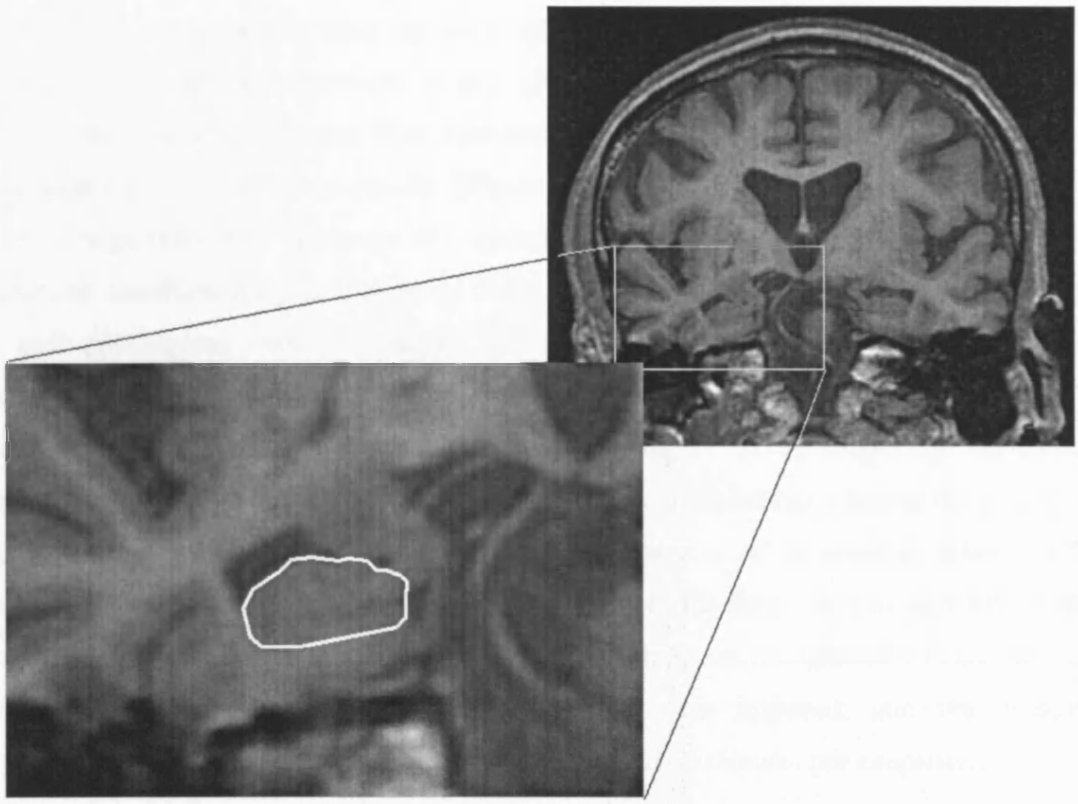


Figure 5.2 Segmentation of the hippocampus.
This structure is outlined coronally and then edited in the sagittal view. The hippocampus includes the hippocampus proper, the subiculum and the alveus.

Prior to ROI measurement, all scans were rigidly registered (six degrees of freedom) to a standard 305 template (Mazziotta *et al.*, 1995) to ensure all scans were in a similar orientation. The repeat images were then accurately registered onto the baseline using a nine degrees of freedom registration (Woods *et al.*, 1998). Scans were measured using mirror-image volumetry (see page 101) with the baseline and registered-repeat scans were presented simultaneously in a random order. ROIs were then manually edited in both coronal and sagittal views. Analysis was performed blinded to the subject's name, diagnosis, left-right orientation of the scan and whether measurements were being performed on the baseline or registered-repeat image. Mean intra-rater variability (calculated as a reliability co-efficient) was 0.98 in the hippocampus and in the cingulate gyrus. Mean intra-rater variability (calculated from the ratio of the absolute difference in measurements to the mean in each person) was 3% in the hippocampus and 4% in the cingulate gyrus. These values were based upon 20 hippocampi segmented twice and six cingulate gyri segmented twice. No time limits were imposed, and the average segmentation time was 45 minutes per hippocampus and 50 minutes per cingulate.

5.2.2.4.Subdivision segmentation

After cingulate segmentation, the region was subdivided into posterior cingulate (PC) which included the retrosplenial cortex, caudal anterior cingulate (CAC) and rostral anterior cingulate (RAC) which approximate to Brodmann areas 23+29+30, 24', and 24 respectively. This was achieved by dividing the cingulate using the anterior and posterior commisures as predetermined cut-offs. Cingulate gyrus anterior to the anterior commisure was labelled rostral anterior cingulate, cingulate posterior to the posterior commisure was labelled posterior cingulate, and the section between was labelled caudal anterior cingulate, see Figure 4.2.

5.2.2.5.Statistical Analysis

Data were analysed using STATA version 8, and SAS (see page 96). Left and right ROI volumes were summed. Volumes for cross-sectional analysis were corrected for TIV with TIV derived according to a previously described protocol (Whitwell *et al.*, 2001). Standardisation was performed separately for the cingulate and hippocampus. For both

structures this was carried out assuming a linear relationship between log-transformed values and log TIV with the slopes of the associations estimated in the control group (Free *et al.*, 1995) (see Equation 4.1). Atrophy rates were analysed on a logarithmic scale [(log (follow-up volume/baseline volume))/interval] in order that doublings and halvings in volume be treated as effects of equal magnitude. Mean atrophy rates were calculated by back transformation with SDs calculated from variance transformation formulae.

Three separate linear mixed models (fitted using proc mixed in SAS) were used to i) compare cingulate rates of atrophy (on a logarithmic scale) in the three patient groups, ii) to compare hippocampal rates with cingulate rates in the three groups and iii) to compare regional cingulate rates of atrophy. Each of these models allowed the variance of, and the covariances between, regional atrophy rates to differ between patient groups. Contrasts of mean levels were used to compare the combined disease groups with controls and to compare the FTLD group with the AD group. In the second model a further contrast was used to investigate whether the magnitude of the difference between cingulate and hippocampal rates differed between FTLD and AD patients. In the third model a further contrast was used to investigate whether the magnitude of the difference between posterior and rostral anterior cingulate rates differed between FTLD and AD patients. Likelihood ratio tests were used to compare variances between groups.

Standard methods were used to calculate sample sizes that gave 90% statistical power to detect a significant difference in atrophy rates between the RAC and PC in AD and FTLD at the 5% significance level.

ROC curves were used to investigate the ability of cingulate and hippocampal rates to differentiate between groups. Their discriminatory ability was compared using logistic regression models relating them individually and together to odds of disease (AD vs. controls and FTLD vs. Controls).

5.2.3. *Results*

Table 5.2 and Figure 5.3 show volumes and rates of atrophy in cingulate and hippocampus in the three subject groups. The mean and standard deviations of baseline and repeat volumes of cingulate and hippocampus show a large degree of overlap between the subject groups.

Table 5.2 Volumes and rates of atrophy of the cingulate gyrus and hippocampus in control, Alzheimer's disease, and frontotemporal lobar degeneration subject groups. All volumes are adjusted for TIV. Rates are expressed as a percentage of baseline volume annualised.

	Controls	**AD**	**FTLD**
Cingulate Volumes (mm^3)			
Mean (SD) - baseline	17100 (2879)	14886 (2325)	14851 (3277)
Mean (SD) - follow-up	17137 (2842)	13653 (2692)	13415 (3357)
Cingulate Atrophy rates (% $year^{-1}$)			
Mean	-0.3	5.9	8.6
95% CIs	-1.1, 0.5	4.2, 7.5	5.1, 11.9
SD	1.2	3.5	4.1
Hippocampus Volumes (mm^3)			
Mean (SD) - baseline	5959 (649)	4672 (567)	3567 (748)
Mean (SD) - follow-up	5984 (623)	4435 (608)	3323 (709)
Hippocampal Atrophy rates (% $year^{-1}$)			
Mean	-0.1	3.4	5.2
95% CIs	-0.7, 0.5	2.3, 4.4	0.6, 9.6
SDs	0.8	2.2	5.4

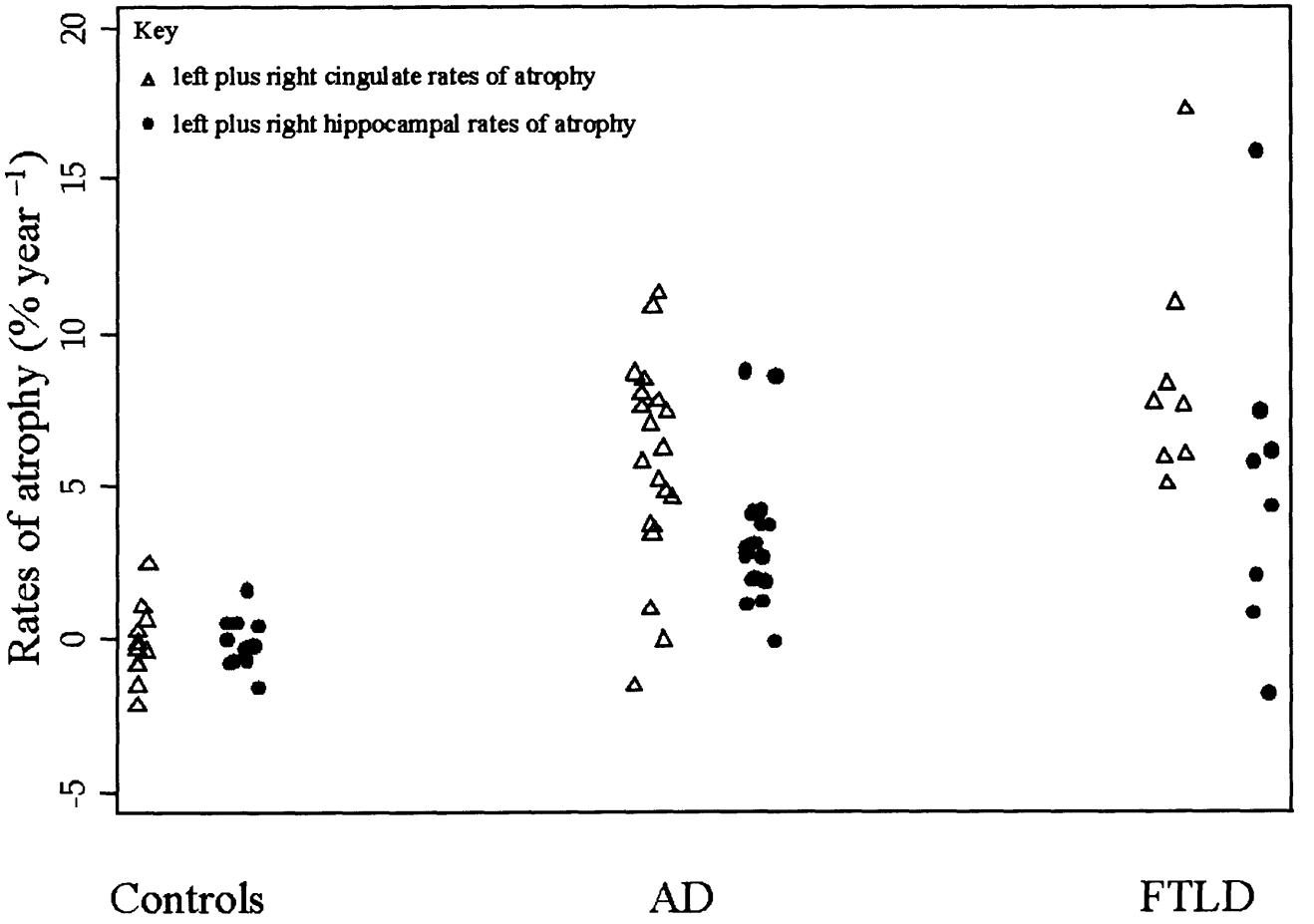


Figure 5.3 Total (left + right) cingulate and total (left + right) hippocampal annualised rates of atrophy in controls, AD and FTLD groups.

Taking the mean of the results in the two disease groups, cingulate atrophy rates were, on average 7.5% per year higher ($p<0.0001$) and more variable ($p=0.001$) than in controls. The mean rate in FTLD patients was slightly higher than that in AD patients, but this difference was not statistically significant ($p=0.7$). Age was not found to be a statistically significant predictor of atrophy rate in the cingulate ($p = 0.6$) and adjusting for age had little effect on the estimated differences between groups.

Again taking the mean of the results in the two disease groups, hippocampal atrophy rates were, on average 4.2% per year higher ($p<0.01$) than in controls. The magnitude of this difference was significantly less ($p=0.01$) than that seen for cingulate atrophy rates. Although hippocampal atrophy rates were somewhat higher in the FTLD patients than in the AD patients this difference was not statistically significant ($p=0.4$). Furthermore there was no evidence ($p=0.7$) that the magnitude of the difference between cingulate and hippocampal atrophy rates differed between the two disease groups. There was no evidence of a significant correlation between cingulate and hippocampal atrophy rates in AD ($p = 0.18$, $R = 0.32$) or FTLD ($p = 0.23$, $R = 0.48$).

The ROC curves presented in Figure 5.4 and Figure 5.5 show the ability of both cingulate and hippocampal rates of atrophy to differentiate between AD and controls and FTLD and controls respectively. Cingulate atrophy rates were significant discriminators of AD from controls ($p=0.01$) although hippocampi were significantly better ($p=0.037$). Cingulate atrophy rates discriminated perfectly between FTLD and controls whereas hippocampal rates were significant predictors ($p=0.044$), although not perfect. Neither cingulate nor hippocampal atrophy rates were significant differentiators between the AD and FTLD groups ($p>0.1$).

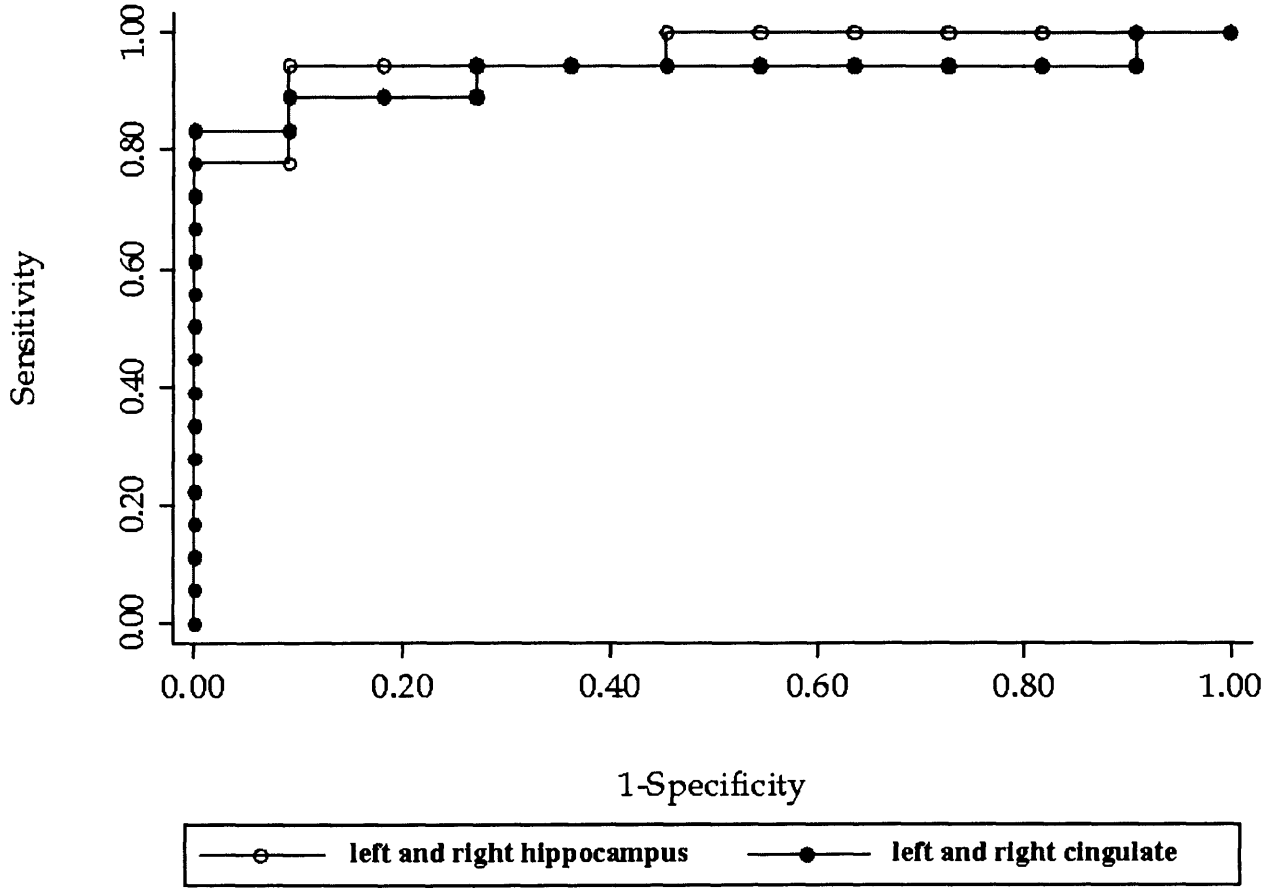


Figure 5.4 ROC curve for the cingulate and hippocampal rates of atrophy in differentiating between AD and controls.

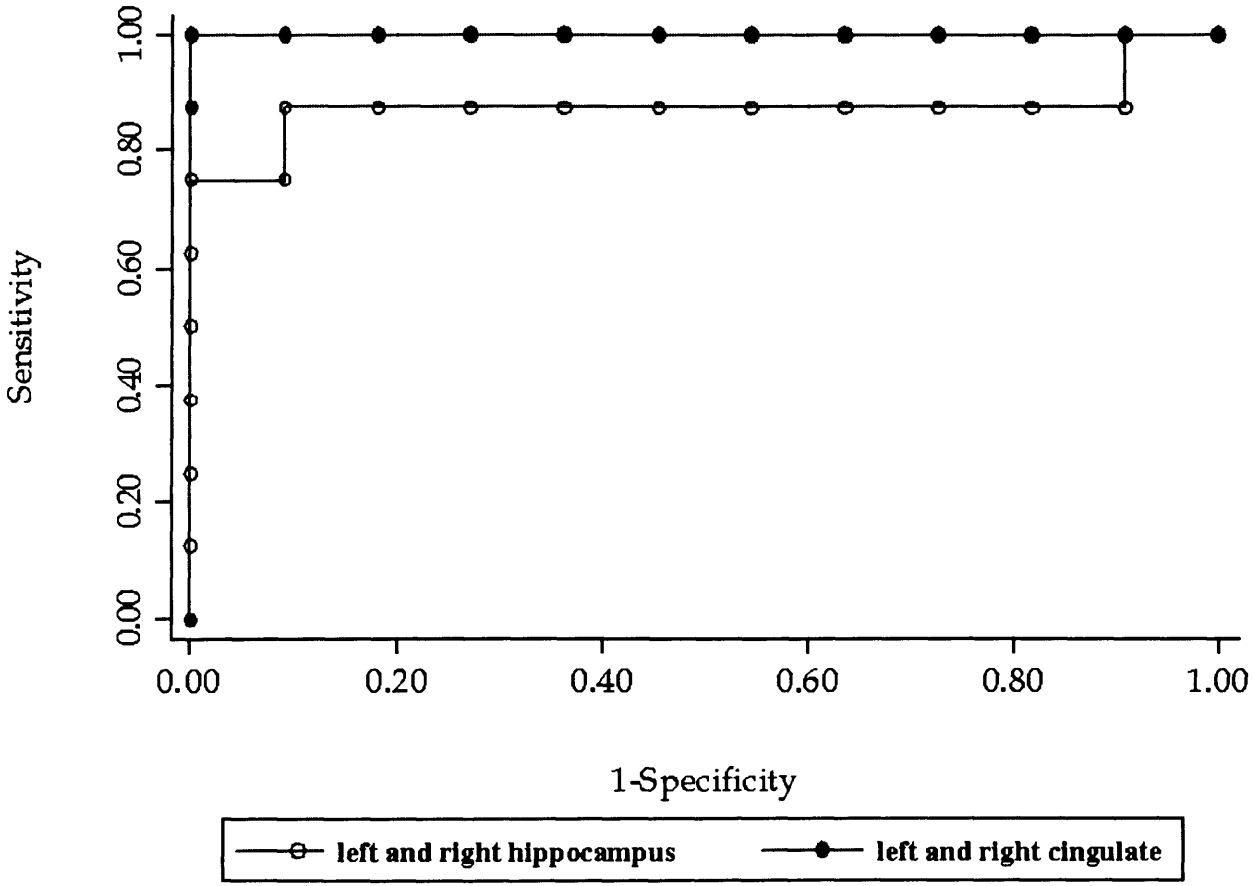


Figure 5.5 ROC curve for the cingulate and hippocampal rates of atrophy in differentiating between FTLD and controls.

5.2.3.1.Subdivisions

Table 5.3 shows mean (SD) rates of atrophy for each subdivision in each subject group. Pooling results across subdivisions and disease groups there was strong evidence ($p<0.0001$) of a difference in mean rates of atrophy between cases and controls. Amongst the AD group rates were highest in the posterior region and lowest in the rostral anterior region whilst among the FTLD this trend was reversed. Neither of these trends achieved statistical significance, and although there was considerable overlap between the two groups there was evidence that the positive and negative trends in the two disease groups were statistically significantly different from each other ($p=0.03$). Retrospective power calculations showed that in order to have 90% power to detect a significant difference in atrophy rates between the PC and the RAC in AD subjects, 79 subjects would be required. Analogous statistics in the FTLD group showed that 131 subjects would be required to detect a difference between the PC and RAC regions.

Table 5.3 Rates of atrophy of subdivisions of the cingulate gyrus in the control, Alzheimer's disease and frontotemporal lobar degeneration subject groups.

Atrophy Rates % year $^{-1}$	**Controls**	**AD**	**FTLD**
Rostral Anterior Cingulate			
Mean	0.1	4.9	9.7
95% CIs	-3.1, 3.3	2.5, 7.3	4.5, 14.6
SD	4.7	5.0	5.8
Caudal Anterior Cingulate			
Mean	0.4	6.4	7.4
95% CIs	-1.4, 2.2	4.5, 8.1	3.8, 10.9
SD	2.7	3.7	4.2
Posterior Cingulate			
Mean	-1.6	7.7	7.6
95% CIs	-3.9, 0.5	4.9, 10.5	-3.8, 11.4
SD	3.3	5.8	4.6

5.2.4. Discussion

This study aimed to establish rates of atrophy of cingulate gyrus and hippocampus in groups of confirmed AD, FTLD and a group of controls. Cross-sectional volumes of hippocampi and cingulate showed a large degree of overlap of the subject groups. Significantly higher rates of cingulate atrophy in both disease groups compared with controls was found. To date there are no other published rates of atrophy of the cingulate in AD or FTLD. However, the main finding of this study is consistent with cross-sectional volumetric data which show decreased volume of the cingulate gyrus in AD (Callen *et al.*, 2001; Killiany *et al.*, 2000).

Significantly higher rates of atrophy for the hippocampus compared with controls was found in both AD (3.4% per year) and FTLD (5.2% per year). Several studies by Jack and colleagues of older sporadic AD cases (mean ages 74-79 years in the different studies) have reported annualised rates of hippocampal atrophy of 3-4% with age-matched controls (mean ages 77-80 years) having rates of atrophy of 1.4-1.7% per year (Jack *et al.*, 1998; Jack *et al.*, 2000; Jack *et al.*, 2004). Other groups have reported rates of atrophy in AD to be 5-6 % per year (mean age ~76 years) and 0.8% per year in control subjects (mean age ~76 years) (Cardenas *et al.*, 2003; Du *et al.*, 2004).

Hippocampal atrophy rates in FTLD were also significantly greater than controls which is consistent with published cross-sectional literature which have shown reduced hippocampal volumes in FTLD (Frisoni *et al.*, 1999; Laakso *et al.*, 2000a). Differences in this study compared with published data may be due to differences in subjects; for example the mean age of the subjects was low and more than half of the AD group had familial AD (Mosconi *et al.*, 2003; O'Riordan *et al.*, 2002). In addition, some differences may be ascribed to variation in segmentation protocols and differences in image quality and acquisition (Pruessner *et al.*, 2000).

These results also show the heterogeneity present in rates of change within the control and disease groups. This heterogeneity may only be partly explained by manual errors, or scan artefact; a significant component of the heterogeneity may be due to true differences in progression of disease, severity of disease or disease duration. With respect to the findings, it is surprising not to find greater heterogeneity within the FTLD group where there was a mixture of pathologies and clinical presentations. Manual outlining errors are inevitable in structures that vary so highly in morphology between subjects and have a number of arbitrary cut offs and subjective judgements in outlining. In particular, it can be difficult to delineate the left and right cingulate gyri in controls with little atrophy as the falx cerebri within the interhemispheric fissure may make the left and right gyri appear continuous.

This study found that hippocampal rates of atrophy were superior to cingulate rates in separating AD from controls although both were significant predictors. Many other imaging methods have been assessed for their ability to separate AD from controls. Within

structural imaging, cross-sectional semi-quantitative measurement of hippocampal atrophy has been shown to discriminate AD from controls with a high sensitivity (87%) and high specificity (93%) (Scheltens *et al.*, 1992). A combination of linear measures of the medial temporal lobe have also proved useful with a sensitivity of 86% for a specificity of 95% (Frisoni *et al.*, 1996b). Cross-sectional quantitative measurement of specific ROIs have also shown similar levels of sensitivity and specificity to this study, especially in the hippocampus and entorhinal cortex (Juottonen *et al.*, 1999; Killiany *et al.*, 2002; Pennanen *et al.*, 2004). Visual analysis and grading of PET images in pathologically confirmed cases gave similar levels of discrimination of AD from controls compared with the volumetric methods reported in this study (Hoffman *et al.*, 2000). Co-varying patterns of cerebral blood flow measured on PET images showed sensitivities of between 76-94% and specificities of 64-81% when separating AD from controls (Scarmeas *et al.*, 2004).

This study found cingulate atrophy to be superior to hippocampal atrophy in separating FTLD from controls. Discrimination was superior to those quoted for one cross-sectional volumetric study (49% sensitivity for the hippocampus and 52% for the entorhinal cortex when specificity was set at 90%) (Frisoni *et al.*, 1999). In another study, change in brain size has shown better separation between FTLD and controls with specificities of around 80% for a sensitivity of 90% (Chan *et al.*, 2001a). However, comparison between FTLD studies is confounded by differences in pathological disease types, ages and disease severities. With respect to separating AD from FTLD, neither cingulate nor hippocampal rates were found to be significant predictors of disease group. This is in keeping with other MR studies which have found cross-sectional volumetric measures of the hippocampus to be poor discriminators of FTLD from AD (Frisoni *et al.*, 1996a; Frisoni *et al.*, 1999). Other imaging techniques have shown good sensitivity and specificity such as SPECT. Comparison of the anterior-posterior gradients of SPET images gave a sensitivity of 86% and a specificity of 93% when separating FTLD from both AD patients and controls (Nagao *et al.*, 2004).

Manual segmentation of ROIs is currently considered to be the "gold standard" in measurement of atrophy in this region. This requires the outlining of the ROI on each MRI "slice". However, errors in manual delineation become very important when differences

are being measured; for example the hippocampus typically loses 3-6 % of its volume over a year in AD (Cardenas *et al.*, 2003; Jack *et al.*, 1998) which is the same order of magnitude as delineation errors in hippocampal measurement (Chan *et al.*, 2001a). As a result, various other methods have been proposed to assess hippocampal size and/or atrophy rates either more easily and/or more precisely (see Chapters 8 and 9: pages 203 and 243) (Crum *et al.*, 2001; Csernansky *et al.*, 2000; Haller *et al.*, 1996). Further research is required to establish whether similar gains can be made in the cingulate region.

A strength of this study is the relative certainty of diagnosis in all patients in the disease groups. However, there was not this level of certainty with respect to the control group apart from one individual who subsequently had a post-mortem revealing a normal brain. One limitation of the study was that scan pairs (baseline and repeat) were sometimes of differing quality. The fact that the mean ages for all groups was young and that the disease groups were predominantly familial cases means this study in not necessarily generalisable to an older, almost exclusively sporadic disease population. The sporadic disease group were relatively young (average age 64 years old compared with 50 in the familial group). This is because the sporadic group were chosen to reduce the confound of age when comparing with the FTLD group and, as a tertiary referral centre, early cases of sporadic AD are often referred. Further research is required in older AD and FTLD groups and in patients with milder disease.

5.2.4.1.Subdivisions

The secondary aim of the study was to establish whether subdivision of the cingulate revealed any differences in atrophy patterns between the disease groups. This study has shown a statistically significant difference in patterns of atrophy along the cingulate between AD and FTLD, with higher rates for AD in the posterior and higher rates for FTLD in the anterior. The higher rate of atrophy in the posterior cingulate in AD is consistent with published data which shows smaller volumes in the posterior cingulate region in AD (Callen *et al.*, 2001; Killiany *et al.*, 2000) or decreased grey matter density in the posterior cingulate (Baron *et al.*, 2001; Scahill *et al.*, 2002). In addition there were higher rates of atrophy in the most anterior portion of the cingulate in FTLD compared with other subdivision rates. Data from current literature is scarce regarding discrete cingulate

subdivision involvement in FTLD although studies have shown predominant involvement of the posterior cingulate (Whitwell *et al.*, 2004a) but potentially reduced posterior cingulate involvement compared with AD (Boxer *et al.*, 2003). The trends shown within groups in cingulate subdivisions, e.g. the posterior section being more affected than anterior in AD and the anterior being more affected than the posterior in FTLD potentially reflecting greater cortical involvement in posterior regions in AD and frontal regions in FTLD, were not significant partly owing to the high variability in these regions. As a result of this great variability, relatively large sample sizes (79 for AD and 131 for FTLD) may be required to detect a significant difference. Further research is required to establish whether better methods of subdivision can be achieved combined with improved atrophy detection or segmentation methodologies.

In conclusion, the cingulate gyrus is severely affected in both AD and FTLD. Rates of cingulate atrophy are as significant as the much more extensively studied hippocampus. Atrophy rates of both structures are significant differentiators of disease groups from controls with similar levels of discriminatory ability implying perhaps that these different components of the limbic system have a similar and possibly concomitant involvement in the AD process. There is a trend for more anterior cingulate atrophy in FTLD and more posterior cingulate atrophy in AD, which may be useful in the differential diagnosis of these conditions.

5.3. Minimum interval needed for hippocampal segmentation

5.3.1. Introduction

Surrogate markers of the atrophy caused by AD are being developed and validated to enable assessment of potentially disease-modifying therapies; the assumption being that alteration in the atrophy rate of the brain would reflect the disease modification. Measures of brain change over time have been shown to be valuable markers of disease progression in untreated patients (Fox, 1999) and as a result have been used as outcome measures in clinical trials (Jack *et al.*, 2003; Krishnan *et al.*, 2003).

Whole brain change may not be as sensitive as localised changes in specific areas of the brain such as the hippocampus. This may be particularly true in the early stages of disease

and it is at this stage of the disease where most benefit may be derived from the patient following their treatment. Therefore areas of the brain that are known to be specifically targeted by the disease process have also been used as outcome measures in clinical trials (Jack *et al.*, 2003; Krishnan *et al.*, 2003).

The interval over which most atrophy rates are calculated is at least one year. This is to ensure that the change over time is greater than the error incurred in estimation of the change. The shorter the inter-scan interval and the more automated the marker, the less resource-consuming a study or trial may be and the faster results can be analysed. In addition, in a clinical setting there may be potential for rapid clinical feed-back to both patient and physician. In this section intervals of less than 12 months were assessed as to whether they can give reasonable rates of atrophy in a group of AD and controls.

5.3.2. Subjects and methods

Subject demographics are described in Table 5.4. Thirty-six subjects from the MIRIAD cohort (see page 92) with probable AD according to NINCDS-ADRDA criteria (McKhann *et al.*, 1984) (see Appendix One), were included. All AD patients had a MR scan compatible with this diagnosis. Twenty age-matched control subjects were also recruited. Ethical approval was received from the Local Research Ethics Committee. All subjects gave written informed consent, and underwent a comprehensive clinical assessment including the MMSE (see Appendix Two). Treatment with an acetyl-cholinesterase inhibitor was not an exclusion criterion for the AD group, however treatment usage was carefully monitored. No control subject had subjective complaints or objective evidence of cognitive impairment. No subject was taking diuretic or steroid therapy during the study.

5.3.2.1.MRI Scan Acquisition

Scans were acquired on a 1.5T Signa Unit (GE Medical Systems, Milwaukee) according to the MIRIAD protocol described in Appendix Three. Each patient was scanned at baseline, six, and 12 months. Scans were performed on the same scanner, using the same acquisition parameters, and each subject was scanned at a consistent time of day.

Table 5.4 Subject demographics.

	Controls	AD
Number	20	36
M/F	10/10	14/22
Mean (SD) age in years	69.1 (7.0)	69.6 (7.3)
Mean (SD) baseline MMSE	29.5 (0.7)	19.4 (4.1)
6m mean interval to baseline (SD) in days	181 (6)	179 (8)
12m mean interval to baseline (SD) in days	364 (18)	365 (18)

AD is Alzheimer's disease, 6m is six month and 12m is twelve month

5.3.2.2.Scan Processing

All scans were transferred to a Sun workstation (Sun Microsystems, Mountain View, CA) for analysis. Regions defining whole brain were obtained (segmented) from all 124 contiguous slices using a semi-automated iterative morphologic technique using MIDAS (see Appendix Four). Each scan triplet (baseline, six, and 12 month) was registered to the MNI 305 brain average (Mazziotta *et al.*, 1995) using a six degrees of freedom (dof) registration. Subsequently, each triplet was co-registered by registering the repeat scans to the baseline image using nine dof. Hippocampi were segmented using the protocol described elsewhere (see Appendix Four) using mirror-image volumetry (see page 95) in

triplets with scans being presented to the operator in a random order with the operator being blinded to order of scans, scan laterality, and subject information.

5.3.2.3.Statistical Analysis

Hippocampal rates of atrophy were calculated for total (left plus right) hippocampus. Data were analysed using STATA version 8 and SAS (see page 96). Annualised rates of atrophy were calculated using log transformed volumes thereby assuming a constant proportionate rate of hippocampal loss. Within each group, differences between annualised atrophy rates calculated using six-month and one-year interval scans were analysed using paired t-tests for means and Pitman's tests for variances. At each time-point, between-group comparisons of atrophy rates were made using t-tests allowing for unequal variances.

Sample sizes for randomised controlled clinical trials required to achieve a particular statistical power to detect a particular proportional reduction in an outcome variable in cases (AD) are proportional to the square of the coefficient of variation (CV) of the outcome variable in cases. The relative numbers of patients required for two different outcome variables is therefore given by the square of the ratio of the respective CVs. 95% bootstrap confidence intervals (bias corrected) for such relative sample sizes (using a logarithmic transform for reasons of symmetry on a relative scale) were calculated using 1000 replicates to compare the two atrophy rates.

5.3.3. Results

Table 5.5 shows the mean (SD) annualised rates of total (left plus right) hippocampal atrophy rates for manual methods using the six and 12 month interval scans in the control and AD groups. Although the mean atrophy rate in AD subjects was slightly higher using a twelve-month interval rather than six, neither this difference (p=0.18) nor the analogous comparison in controls (p=0.69) was significant. However, at an individual level there were considerable differences in rates measured over the different intervals (Figure 5.6 A: 95% reference range on the difference: -4.24 to 4.96). Results from six-month interval scans were more variable than that at 12-months in controls (p=0.007) (Figure 5.6 B). In AD subjects the difference was not significant (p=0.98), perhaps reflecting the lower power

to detect differences through greater between-subject variability in atrophy rates. Using AD subjects, the relative sample size for the atrophy rates at 12 months compared with six did not show a significant reduction in sample size at 12 months 0.71 (95% CIs 0.29 1.50).

Table 5.5 Annualised atrophy rates calculated as a percentage of baseline volume from the six and 12 month interval scan data using manual measures.

	Controls	**AD**	**p value**
Annualised mean (SD) 6 month rate	0.41 (1.69)	3.95 (3.01)	< 0.0001
Annualised mean (SD) 12 month rate	0.28 (0.93)	4.57 (2.98)	<0.0001

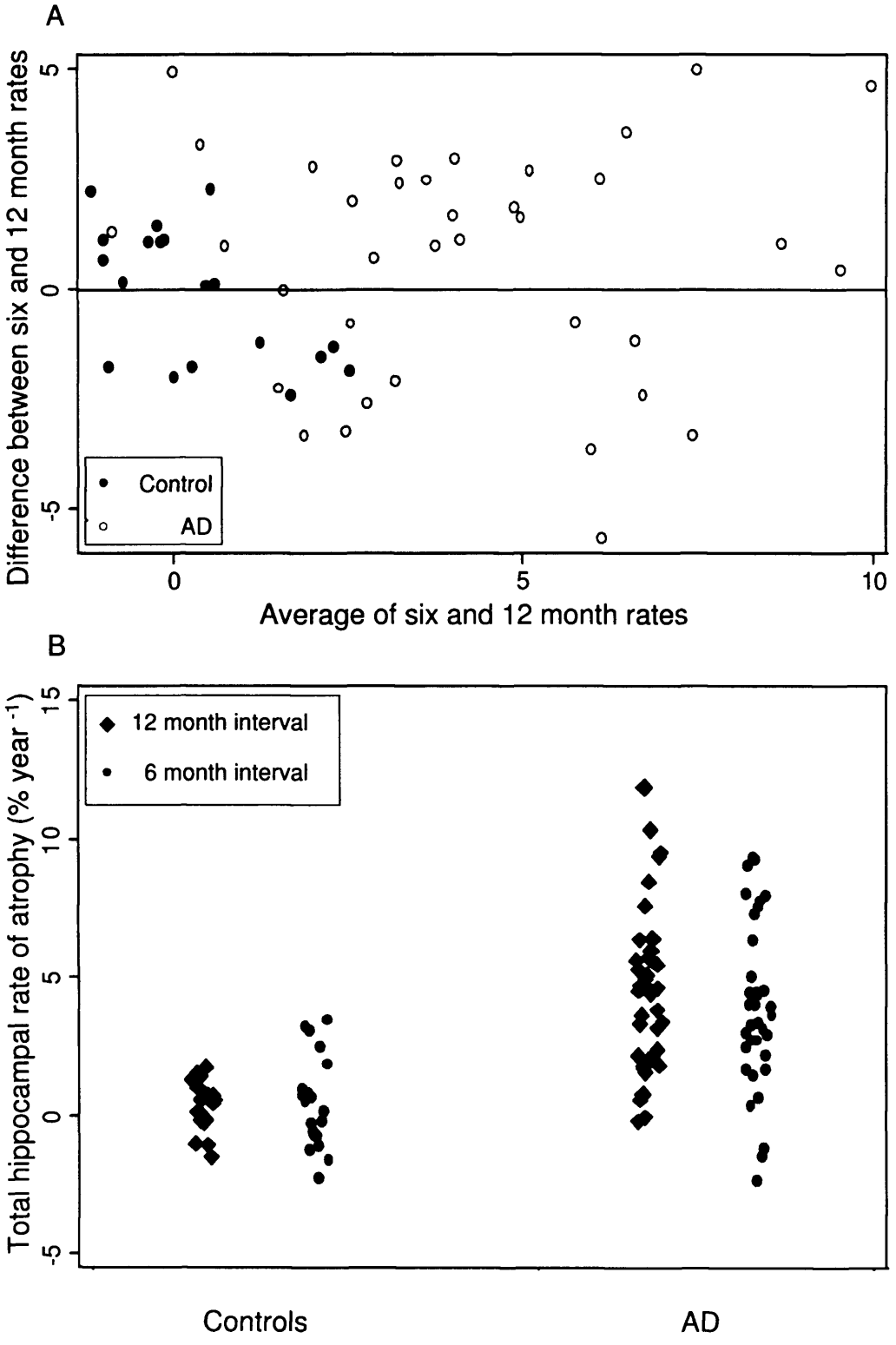


Figure 5.6 A) Bland-Altman plot of six and 12 month interval atrophy rates, B) annualised atrophy rates expressed as a percentage of baseline volume using manual methods.

5.3.4. *Discussion*

In this study it was found that annualised atrophy rates from six-month interval scans were similar to twelve-month interval scans. The variances were higher using the six-month interval scans, which is not surprising since any errors would have been doubled by annualising the rates. However, the increase in variance was not sufficient to make the differences in sample sizes calculated between the six- and twelve-month interval rates significantly different. The mean rates of atrophy in the AD and controls groups were also slightly but non-significantly lower using the six-month interval than the twelve-month interval rates in both controls and AD patients. This may be simply chance or because the amount of atrophy seen between the six month and baseline scans was not large enough to visualise and therefore these differences are not incorporated into the segmentation of the hippocampus. The effect of this being a potential bias was theoretically reduced by being blinded to the date and loading the scans in a random rather than chronological order. However, in some cases the order of the scans may be obvious to the operator, especially when a large amount of atrophy has occurred.

Mean atrophy rates were around 3-4.5% per year in AD compared with less than 0.5% per year in controls. This is within the range of previously published hippocampal rates of atrophy for AD patients of this age group; 3-4% per year (Jack *et al.*, 1998; Jack *et al.*, 2000; Jack *et al.*, 2004) 5.4% per year (Cardenas *et al.*, 2003) 5.9% per year (Du *et al.*, 2004). However, the rate of atrophy in the control group is lower than that quoted in these studies; 0.8% per year (Du *et al.*, 2004) 1.4-1-9% per year (Jack *et al.*, 1998; Jack *et al.*, 2000; Jack *et al.*, 2004) 1.8% per year (Cardenas *et al.*, 2003) 2% per year (Kaye *et al.*, 1997). Differences in the ages of these control groups may account for some of the variation in reported rates of atrophy (Scahill *et al.*, 2003). Additionally, the control subjects in this study may have been healthier. Other explanations of differences, for example the use of different segmentation methodologies or scanning protocols by different research groups, may also account for some of the differences reported (Pruessner *et al.*, 2000).

Despite finding large and significant differences in AD subjects' rates of atrophy from controls a large degree of overlap between subject groups was still observed. This may reflect true group heterogeneity, scan artefacts or measurement error in both subject groups. In the AD group this may reflect different progression or stage of disease. In this study there was histological (autopsy) confirmation of disease in only two of the AD group and one of the controls, therefore there may be some uncertainty as to the pathological status of all of the subjects. One strength of this study is that one operator (myself) carefully segmented all hippocampi. This means that inter-rater variability was not a confounding factor in the calculated manual atrophy rates.

Only one other study of hippocampal atrophy over short inter-scan interval imaging has been reported in the literature. However in this study, with an inter-scan interval of 24 weeks, the amount of hippocampal loss was not reported as an annualised rate of atrophy it was instead calculated as a total percentage reduction (Krishnan *et al.*, 2003). The percentage reduction over this period of time was calculated to be 8.2% in the placebo arm, leading to an approximate and surprisingly high annualised rate of 18% loss per annum. This rate exceeds all others quoted in the literature (see above and Table 1.3). Other short interval studies have been performed measuring whole brain change and change of ventricle volume. One study looked at whole brain change and whole ventricular change in short interval imaging using multiple scanning time-points (Bradley *et al.*, 2002). This small study (seven AD subjects) found that increase in ventricular volume and decrease in brain volume over short inter-scan intervals was a reliable marker of possible or probable AD. This study also established that increasing the inter-scan interval to one year decreased the number of people needed for a trial required to detect a reduction in atrophy rate although this was not significant. The greater 12 month interval also reduced the amount of people required to detect a standard amount of atrophy owing to increased precision.

Other published work using data from this study found that ventricular change, measured using ventricular BSI, was a good marker of disease status using a six month interval (Schott *et al.*, 2005). This study also showed that increasing the interval to one year decreased the number of patients required as a result of increased precision. However, it may be that localised changes warrant assessment, and with increasingly automated

techniques available for measuring hippocampal change, this may be possible as an additional measure with little extra operator input. Reductions in scan acquisition artefact may also improve the utility of hippocampal measures.

5.3.5. Conclusions

Change in the hippocampus can be measured at intervals less than one year; however there is as expected an increased variance in annualised atrophy rates with the shorter scanning intervals. Nonetheless, rates over six months were consistent with the 12 months results at a group level. This finding has implications in both clinical and trial settings where a six month interval may allude to the eventual results and allow more frequent monitoring of patients.

5.4. Chapter conclusions

Substructures of the brain can be measured precisely enough that change can be calculated over time. The amount of change and its location has been shown to be different in different diseases (AD vs. FTLD in both the cingulate and hippocampus). Application of manual delineation to serial imaging reveals patterns of atrophy which can be used to differentiate different diseases. Subdivision of structures is possible and can reveal underlying patterns of atrophy that may differ between different groups in a large structure which is affected in both AD and FTLD. Manual measures can be precise enough to measure changes at intervals less than one year. These results have been shown in well-defined groups of patients, many of which contain subjects which have pathological- or genetic-confirmation of their diagnosis. This work gives a benchmark with which more automated techniques can be compared.

6. COMBINED CROSS-SECTIONAL AND LONGITUDINAL STUDIES

6.1. Chapter Introduction

Background

As shown in Chapters 4 and 5, both cross-sectional and longitudinal studies can further the understanding of degenerating diseases and aid their diagnoses. Cross-sectional studies give a measure of the amount of loss before the scan was taken and the existing normal variation. Longitudinal studies, which calculate the rate of loss in a structure give a measure of the progression of the disease and have less confounding problems with inter-individual variability as each person acts as their own control. Owing to the different information provided by each approach, studies which combine information from both are often useful when investigating disease-related questions.

Chapter outline

In this chapter, results are presented from two studies which combine both cross-sectional and longitudinal data to answer specific questions: i) whether hippocampal asymmetry is a potential marker of AD and ii) whether hippocampal atrophy can be detected in presymtomatic cases of familial AD.

6.2. Is hippocampal asymmetry affected by AD?

6.2.1. Introduction

There is considerable interest in using MRI measures as an aid to early diagnosis. Numerous imaging studies have shown that cross-sectional hippocampal atrophy is a consistent finding in AD (Chan *et al.*, 2001b; Jack *et al.*, 1997; Lehtovirta *et al.*, 1995); indeed the presence of hippocampal atrophy has been shown to be predictive of a diagnosis of AD and of symptom onset (de Leon *et al.*, 1989; Fox *et al.*, 1996b; Kaye *et al.*, 1997; Rusinek *et al.*, 2003). There have been some longitudinal MRI studies of hippocampal volumes, and these have shown that rates of hippocampal atrophy in patients with AD are significantly increased compared with controls (Du *et al.*, 2004; Jack *et al.*, 1998; Jack *et al.*, 2000; Jack *et al.*, 2004; Laakso *et al.*, 2000b).

In contrast to focal neurodegenerative conditions such as FTLD, AD is usually considered to involve generalised symmetrical cortical atrophy (Chan *et al.*, 2001b). However, the normal human brain has some asymmetry. Right-greater-than-left asymmetry has been reported in normal subjects when whole hemisphere (Gur *et al.*, 1991), temporal (DeCarli *et al.*, 1994; Jack *et al.*, 1989) or frontal lobes (Howard *et al.*, 1995; Weinberger *et al.*, 1981) are compared. Several MRI studies have shown the right hippocampus to be larger than the left in normal subjects (Csernansky *et al.*, 1998; Csernansky *et al.*, 2000; Geroldi *et al.*, 2000; Jack *et al.*, 1989; Laakso *et al.*, 2000b; Mervaala *et al.*, 2000; Pruessner *et al.*, 2000; Soininen *et al.*, 1994; Wang *et al.*, 2001), with estimates of the magnitude of this asymmetry of up to 10% (Soininen *et al.*, 1994). However, not all studies report significant hippocampal asymmetry (Bhatia *et al.*, 1993; Fox *et al.*, 1996b; Golebiowski *et al.*, 1999; Kaye *et al.*, 1997) and two studies have reported the left hippocampus to be marginally larger than the right (Ashtari *et al.*, 1999; Cook *et al.*, 1992).

The effect of the progression of AD on hippocampal asymmetry is unclear. Whilst one study suggested that the natural right-greater-than left asymmetry in brain structures is reduced or even reversed in AD (Geroldi *et al.*, 2000), another reported considerable heterogeneity in the degree and direction of hippocampal asymmetry in a very small number of patients with early AD (Fox *et al.*, 1996b). Two studies have reported a dose-dependent reduction in the degree of hippocampal asymmetry associated with the APOEε4 allele (a known genetic risk factor for AD), with homozygotes exhibiting reversal of the normal asymmetry (Geroldi *et al.*, 2000; Soininen *et al.*, 1995), although another study failed to reproduce this finding in a group of cognitively normal subjects homozygous for APOEε4 (Reiman *et al.*, 1998). Soininen *et al.* (Soininen *et al.*, 1994) demonstrated a reduction in the degree of hippocampal asymmetry in subjects with age-associated memory impairment. If hippocampal asymmetry does alter in early AD the implication is that asymmetric atrophy must be occurring at an early stage of the disease, suggesting that one hemisphere or hippocampus is more vulnerable to the disease.

In addition to being of possible diagnostic value, it is important to understand the natural history of hippocampal changes in AD since measures of hippocampal volume have been proposed as outcome measures in trials of potentially disease-modifying therapies in AD.

It is important to understand whether there is a left-right difference as it might be possible to confine measurement of this structure (which is operator-intensive) to one hippocampus. The aim of this study was to establish whether there is true asymmetry in hippocampal volumes or atrophy rates in both patients and controls, using a technique (mirror-image volumetry) in which the investigator is blinded to the right-left orientation of the scans, thereby preventing laterality bias.

6.2.2. *Materials and Methods*

6.2.2.1.Subjects

Subjects were seen at the Dementia Research Centre at The National Hospital for Neurology and Neurosurgery, and were all taking part in ongoing neuroimaging studies. Ethical approval for the study had been given by the Local Research Ethics Committee, and all subjects gave written informed consent. All subjects underwent clinical assessment including the MMSE (Folstein *et al.*, 1975) (see Appendix Two), and all patients in the AD group fulfilled standard NINCDS-ADRDA criteria (McKhann *et al.*, 1984) for the diagnosis of probable AD (see Appendix One).

Fifty controls and 32 patients with AD each had two serial volumetric MRI assessments. All patients in this study had their MRI scans reported by an experienced neuroradiologist as compatible with the diagnosis of AD. The AD group comprised 13 FAD and 19 SAD cases. Genetic confirmation of disease was available in 11 FAD cases and one post-mortem confirmation of disease was available on a further FAD case. APOE genotyping was available on a subset of AD patients (13 FAD and 11 SAD). Subject demographics are shown in Table 6.1.

6.2.2.2.Imaging

T1-weighted volumetric MR brain scans were performed on a 1.5 Tesla Signa unit (General Electric, Milwaukee), using the standard protocol for all subjects except 11 AD and 11 controls who were scanned using the MIRIAD protocol (see Appendix Three). All scans were transferred to a Sun workstation (Sun Microsystems Inc., Mountain View, CA) for analysis. The software package MIDAS (Freeborough *et al.*, 1997) was used for all manual segmentation. Each hippocampus was carefully manually edited in both the coronal and sagittal planes to improve accuracy of delineation (see Appendix Four).

Table 6.1 Subject demographics.

	Controls	AD
Number	50	32
M/F	26 / 24	13 / 19
Age (Years)	59.6 (13.8)	59.0 (11.3)
Mean (SD) baseline MMSE (Max=30)	29.4 (0.8)	19.4 (5.0)
Mean interval (Days)	418 (172)	456 (311)
Sporadic / Familial	NA	19 / 13
APOE status	Not available	24 APOE (11-/-, 12 -/ε4, 1 ε4/ ε4, 8 not available)

NA = not applicable

6.2.2.3.Cross-sectional and Longitudinal Measurements

Prior to hippocampal measurement, the follow-up scans were accurately registered onto the baseline images, using a nine degrees of freedom (dof) registration algorithm (Freeborough *et al.*, 1996b). The hippocampus was always measured on the right-hand side of the presented image using mirror-image volumetry (see page 101). The baseline and registered repeat scans were presented simultaneously in a random order. The first hippocampal region was pasted onto the second scan to ensure consistency of arbitrary boundaries

between scans. This region was then manually edited where necessary. The investigator was blinded to the subject's name, diagnosis, left-right orientation of the scan and whether measurements were being performed on the baseline or registered-repeat image.

6.2.2.4.Statistical Analysis

Data were analysed using STATA version 8 (see Chapter 3, page 96). Geometric means of the right and left hippocampal volumes and their ratio were calculated according to disease status and visit. A student's one-sample t test was used to investigate whether there was significant asymmetry in either the hippocampal volumes, or the atrophy rates in the two groups. Two sample t-tests (assuming unequal variances) were used to make comparisons between the disease and control groups. F tests were used to investigate the difference in standard deviation in ratios and raw hippocampal volumes between the two groups. Standard methods were used to calculate retrospective sample sizes that gave 90% statistical power to detect a significant difference in atrophy rates between the right and left hippocampal rates in AD at the 5% significance level. Similar methods were used to calculate the sample sizes required to detect a significant difference in atrophy rates between the APOE ε4+ and APOE ε4- carriers.

6.2.3. 6.2.3 Results

Table 6.2 shows geometric mean levels of right and left hippocampal volumes and their ratio in AD patients and controls at baseline and repeat visit. Comparison of the changes in this ratio over time provided evidence (p=0.02) of a difference between groups as shown in Figure 6.1. In controls the right hippocampal volumes were, on average, around 2% (mean of both visits 1.7% CI –0.3 and 3.7%) greater than the left hippocampal volumes. This was not statistically significant (p=0.1). In AD cases, at baseline, right hippocampal volumes were also, on average, about 2% greater than left hippocampal volumes. However at the follow-up visit (on average 15 months later) right and left hippocampal volumes were similar. This reduction in the right to left ratio in cases was statistically significant (p=0.02). Additionally, the variability in the ratio of one volume to the other was substantially larger in the AD cases than in controls at baseline (p=0.02) but not at repeat (p=0.06). Figure 6.2 shows the hippocampal ratio of ratios (right/left hippocampus at baseline divided by the same measure at repeat imaging). This figure demonstrates the significant decrease in right/left asymmetry in AD subjects and the greater variability in this

measure over time compared with the control group. There was no evidence of differences in hippocampal ratio by gender at baseline or repeat imaging ($p > 0.1$). There was no evidence of a significant association between hippocampal ratio at baseline and either MMSE ($p = 0.84$), or age ($p = 0.32$) in the AD group. There was no evidence of a significant association between age and hippocampal ratio at baseline imaging in controls ($p = 0.52$).

The reduction in the right to left ratio in AD patients was reflected in greater atrophy in the right than the left hippocampus although this just failed to reach statistical significance ($p=0.05$). In controls, atrophy rates were much lower than in AD patients ($p<0.001$, both for left and right) and similar in the two hippocampi ($p=0.9$). Retrospective power calculations showed that in order to detect a significant difference in atrophy rates between the right and left hippocampus in AD subjects, 115 subjects would be required.

The presence of an APOEε4 allele had no significant effect on left or right hippocampal size or their ratio at baseline and repeat in this study ($p>0.4$). Mean (95% CI) right/left hippocampal ratios at baseline for the APOEε4– and APOEε4+ groups were 1.01 (0.92, 1.10) and 1.03 (0.97, 1.09) respectively ($p=0.76$). At repeat the mean (95% CI) ratios for the APOEε4– and APOEε4+ groups were 0.99 (0.92, 1.06) and 1.00 (0.94, 1.06) respectively ($p=0.80$). The presence of an APOEε4 allele had no significant effect on the atrophy rates of either the left or the right hippocampus in this study (both $p>0.2$). Mean (95% CI) rates for the left and right hippocampus in the APOEε4– patients were 3.3 (1.1, 5.4) and 5.3 (2.9, 7.8) respectively; in the APOEε4+ patients these were 4.9 (2.4 7.4) and 6.9 (4.7 9.2) respectively. Retrospective power calculations revealed that in order to detect a difference in left hippocampal rates between APOE ε4+ carriers and non-carriers, 105 subjects would be required. Analogous statistics for the right hippocampus revealed that 122 subjects would be required to show a difference.

Table 6.2 Mean (95% CI) hippocampal volumes at baseline and repeat imaging with corresponding atrophy rates in control and AD groups.

		RIGHT (mm^3)	LEFT (mm^3)	R/L Ratio	ATROPHY (%/YR)	
					RIGHT	LEFT
CONTROLS	BASELINE	2715 (2631, 2801)	2669 (2586, 2754)	1.017 (0.996, 1.039)	1.1 (0.5, 1.8)	1.2 (0.5, 1.8)
	REPEAT	2683 (2600, 2769)	2637 (2551, 2727)	1.017 (0.996, 1.039)		
AD	BASELINE	2262 (2110, 2425)	2220 (2075, 2375)	1.019 (0.981, 1.058)	6.3 (4.9, 7.8)	4.6 (3.3, 6.0)
	REPEAT	2095 (1941, 2261)	2108 (1970, 2256)	0.994 (0.957, 1.032)		

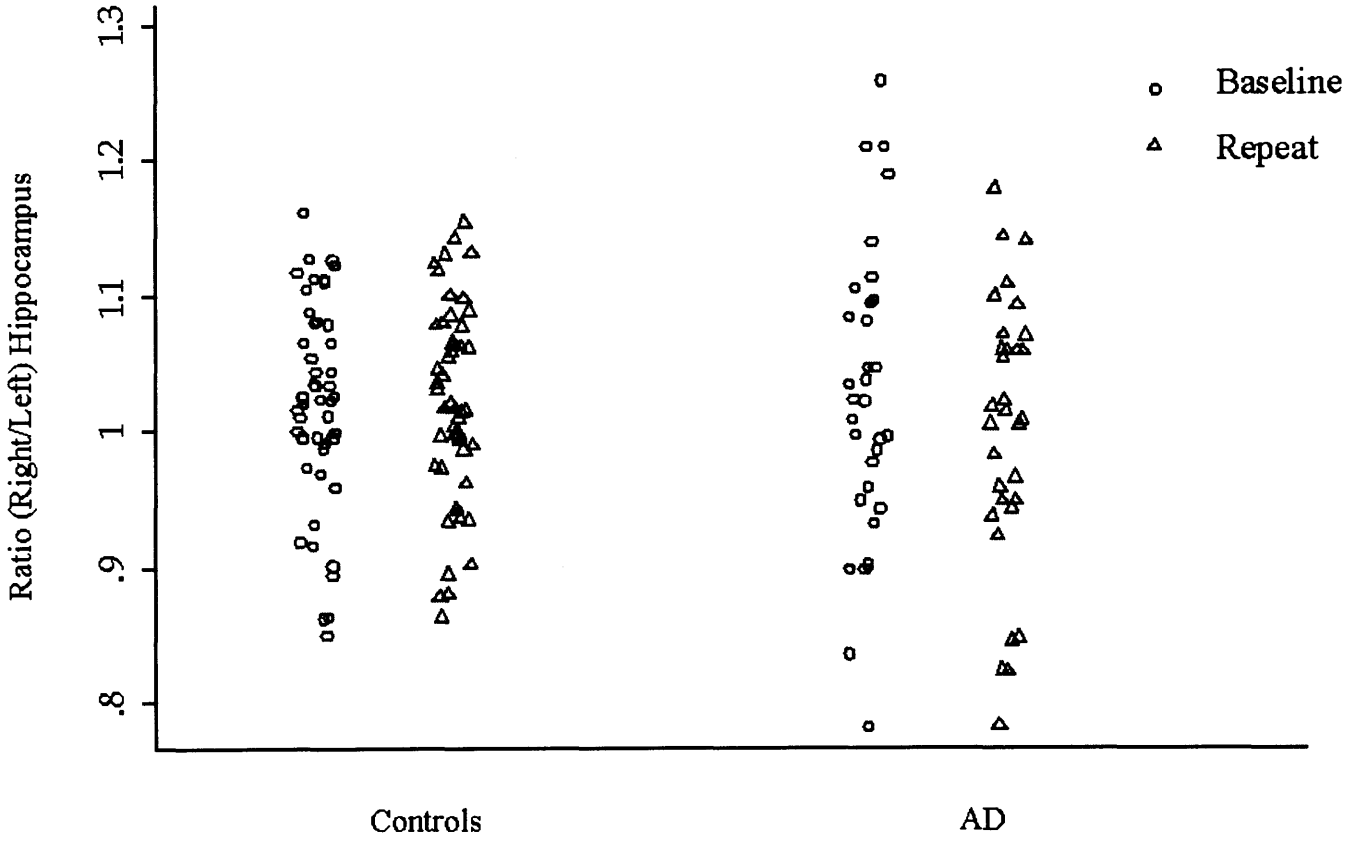


Figure 6.1 Ratios of hippocampi (right / left) at baseline and repeat imaging in control and AD groups.

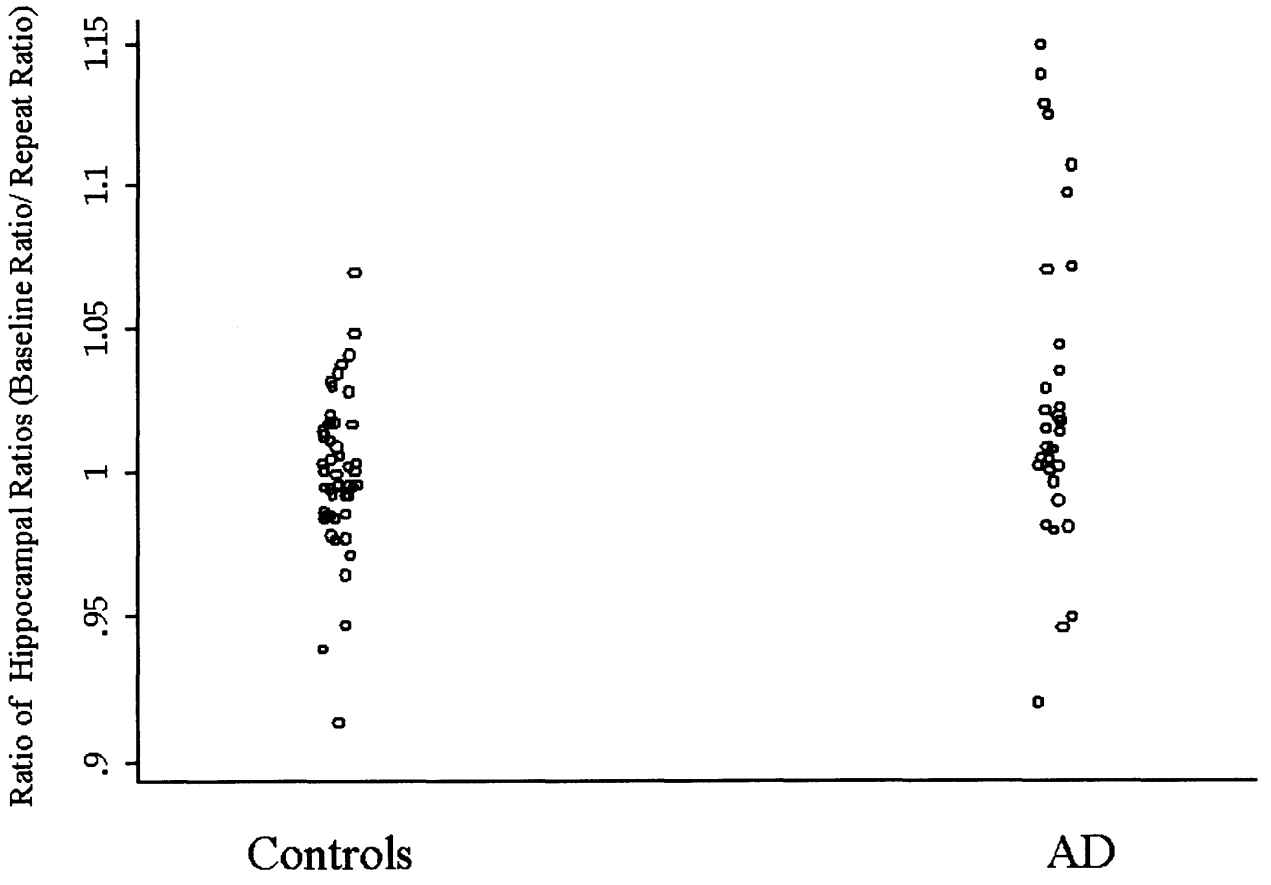


Figure 6.2 Ratio of hippocampal ratios in control and AD groups. (Right/left hippocampal ratio at baseline / right/left hippocampal ratio at repeat imaging).

6.2.4. *Discussion*

Normal Controls

In this study, a trend towards right-greater-than-left hippocampal volume asymmetry in healthy controls was demonstrated, which concurs with previous studies (Howard *et al.*, 1995; Jack *et al.*, 1989). An average right>left asymmetry of approximately 2% in normal controls was reported, which is somewhat lower than previous studies (Geroldi *et al.*, 2000; Soininen *et al.*, 1994). One possible cause of this discrepancy might be the different segmentation methodologies employed by different research groups (Pruessner *et al.*, 2000). In particular, there may be variations in the boundaries chosen for hippocampal delineation. In this study, to aid consistency of the arbitrary boundaries of the segmentation both baseline and repeat images were presented to the rater simultaneously. In addition, raters were blinded to chronological order and laterality of the scans and patient information to limit potential measurement bias. Another possible explanation for the difference between this and other studies may result from differences in the composition of groups selected for such studies. Of two studies assessing healthy elderly cohorts, one reported no significant hippocampal asymmetry (Kaye *et al.*, 1997), whilst another found a minor left>right asymmetry (Ashtari *et al.*, 1999), raising the possibility that hippocampal asymmetry may be affected by age, or that asymmetry is an inconsistent finding across different subject groups.

Alzheimer's Disease

The hippocampal volume reductions in individuals with AD were consistent with numerous previous publications (Convit *et al.*, 1993; Du *et al.*, 2004; Kohler *et al.*, 1998; Krasuski *et al.*, 1998; Laakso *et al.*, 1995; Mega *et al.*, 2002).

Several cross-sectional studies have suggested that alterations in right-greater-than-left asymmetry in hippocampal volume may be associated with age-associated memory impairment (Soininen *et al.*, 1994), the possession of the APOEε4 allele (Geroldi *et al.*, 2000; Soininen *et al.*, 1995), or the process of AD (Geroldi *et al.*, 2000), and that such changes reflect altering rates of left/right hippocampal atrophy as the disease progresses. However, other studies have suggested that AD may be associated with preferential involvement of the left hemisphere (Scahill *et al.*, 2002; Thompson *et al.*, 2001; Thompson *et al.*, 1998). Metabolic deficits have also been shown to be more severe on the left than the right in individuals with increased genetic risk of developing AD (Small

et al., 2000). Another study suggested there was no interaction between laterality of the hippocampus and AD – control volume differences (Du *et al.*, 2004).

A non-significant right>left asymmetry of approximately 2% was found in the subjects with AD at baseline but no asymmetry at repeat. This was caused by a borderline-significant trend for higher atrophy rates in the right hippocampus. This resulted in a decline in the degree of asymmetry in the patient group over the period of assessment which was just statistically significant, however, as Figure 6.1 shows there is a wide range in the right/left ratios. Retrospective sample size calculations showed that a sample of 115 subjects would be required to detect a significant difference between left and right hippocampal rates in the AD group.

Hippocampal atrophy rates were greater in patients than controls, in keeping with previous longitudinal studies (Du *et al.*, 2004; Jack *et al.*, 1998; Jack *et al.*, 2000; Jack *et al.*, 2004; Laakso *et al.*, 2000b) (see Table 1.3). The combined hippocampal atrophy rate of 5.5% per annum in patients with AD (including those with young onset FAD) is similar to other studies (3.5% (Jack *et al.*, 2000), 4.0% (Jack *et al.*, 1998), 3.0-3.6% (Jack *et al.*, 2004) 5.4% (Cardenas *et al.*, 2003), 5.9% (Du *et al.*, 2004), 7.2% (Laakso *et al.*, 2000b), 4-8% (Fox *et al.*, 1996b)). If hippocampal asymmetry changes over the course of AD, as suggested by Geroldi *et al.*, a significant right-greater-than-left atrophy rate should be detected by longitudinal studies. In previous longitudinal series (28 subjects with AD), Jack *et al.* concluded that annual rates of volume change for the left and right hippocampus did not differ (Jack *et al.*, 1998). Laakso *et al.* found significant loss of right hippocampus compared to controls, but non-significant loss of left hippocampus compared to controls (Laakso *et al.*, 2000b) suggesting the possibility of greater right hippocampal atrophy.

In this study, the fact that nearly half of the patients had familial rather than sporadic AD might have independently affected hippocampal atrophy rates. Increased variability within the AD atrophy rates compared with controls was also found which may be a reflection of a variety of disease severities or true heterogeneity. Greater variability in the AD group is supported by the finding of increased cortical variability (variability of deep sulci) in subjects with AD compared to controls (Thompson *et al.*, 1998). This increased variability may possibly explain the range of results in the literature and

reduces the power of a relatively small study to comment on extremely subtle volumetric findings.

No significant association between the asymmetry at baseline or repeat visit and the presence of the APOEε4 allele was found. This is consistent with a number of studies (Jack, Jr. *et al.*, 1998; Reiman *et al.*, 1998). In addition, it was found that the APOEε4 allele had no significant effect on the atrophy rates of either the left or right hippocampus. These findings are consistent with a number of reports (Bigler *et al.*, 2002; Jack *et al.*, 1998) but contradict findings of another study which suggested a gene dose effect of APOEε4 causing increased reversal of asymmetry (Geroldi *et al.*, 2000). This study, much like the results reported in this chapter, had relatively little power to detect subtle effects of numbers of patients 14 APOEε4 -/- 9 APOEε4 +/- and 5 APOEε4 +/+. The study reported in this section of this chapter had relatively small numbers of patients with APOE confirmed status although one would expect if there were any effects of the magnitude found by Geroldi then a significant trend in either asymmetry or atrophy rates would have been found. However, it must be noted that the presence of at least one APOEε4 led to higher mean atrophy rates in both left and right hippocampi in the results reported in this chapter. It may be that owing to the high variability of, and small numbers in, the two groups, there was not enough power to detect a significant difference. Retrospective sample size calculations showed that a sample of 115 subjects would be required to detect a significant difference between left and right hippocampal rates in the AD group. In addition, as nearly half the AD group consisted of familial cases it is possible that any subtle effects caused by the ε4 allele were masked. Further research is required to establish any interaction between the ε4 allele and other known genotypes in these familial cases.

Whilst these findings suggest a disease-related reduction in asymmetry caused by increased right hippocampal atrophy a greater variability in asymmetry and atrophy rates is also reported. For these reasons, it is suggested that alterations in the normal right-left asymmetry of the hippocampus are not necessarily a reliable indicator of disease progression in AD. Furthermore, caution is suggested in accepting the right hippocampus is a superior surrogate marker of AD compared to left hippocampus or that hippocampal measurement should necessarily be restricted to the right in studies. Instead, taking the average of left and right hippocampal measures may reduce the effect of measurement error without a loss of sensitivity to change.

6.2.5. *Conclusions*

Using mirror-image volumetry it has been demonstrated that a non-significant right>left asymmetry in hippocampal volumes in healthy subjects. Right>left asymmetry was present in the AD at baseline but reduced at repeat (approximately 15 months later) and this measure was associated with greater variability compared to controls. Reduction in asymmetry was caused by a borderline significantly greater atrophy rate in the right hippocampus. In this group, neither volumes, asymmetry ratios nor rates were affected by the presence of an ϵ4 allele.

6.3. *Hippocampal volumes and atrophy rates in presymptomatic FAD*

6.3.1. *Introduction*

The hippocampus is one of the earliest sites of pathological alteration in AD (Braak *et al.*, 1993). The pathology has been shown to be present in about 85% of subjects with MCI, a stage considered to be a transitional between normal ageing and AD (Morris *et al.*, 2001). Quantification of AD pathological changes within the hippocampus have been shown to correlate with hippocampal volumetric measures using post-mortem MRI (Bobinski *et al.*, 2000; Gosche *et al.*, 2002). In addition, several *in-vivo* imaging studies have shown significant hippocampal volume reduction in patients with a clinical diagnosis of AD compared with controls (Jack, Jr. *et al.*, 1992), even in the early mild stages of the disease (Jack *et al.*, 1997) and in patients with MCI (Jack *et al.*, 1999).

Although most AD cases are sporadic, rare familial forms of the disease exist with an autosomal dominant pattern of inheritance. As discussed in Chapter 1, page 20), mutations in three genes have been identified; APP, PS-1 and PS-2 (Hardy, 1997). FAD results in significantly younger ages of onset than SAD; however, patients develop a similar pattern of insidious onset and progressive decline affecting episodic memory followed by global cognitive dysfunction (Fox *et al.*, 1998).

Serial imaging of mutation carriers from presymptomatic to affected provides an opportunity to track MRI changes that may precede clinical manifestations. Their relatively young age at disease onset minimises the confounding effect of age-related co-morbidity. In a previous longitudinal MRI study from this centre, Schott *et al.*

estimated that atrophy of medial temporal lobe structures preceded the clinical onset of the disease by an average of 3.5 years (Schott *et al.*, 2003). This was, however, based on only 5 FAD subjects who had 15 scan points in total over 2-4 years of assessment.

This study expands on previous work by including further FAD subjects and scanning time-points. This permits the use of hierarchical modelling of how hippocampal and whole brain volumes change as the disease progresses from the pre-symptomatic stage through to AD diagnosis.

6.3.2. Methods

6.3.2.1.Subjects

Nine carriers (four males) of autosomal dominant AD mutations (five with mutations in APP and four with mutations in PS-1) were recruited from the Cognitive Disorder's Clinic at the National Hospital for Neurology and Neurosurgery. Twenty-five age- and gender-matched controls (2-3 controls per mutation carrier) were recruited from spouses (n=7), unaffected relatives including those with negative genetic test results (n=4), and healthy volunteers (n=14). All subjects gave informed written consent with assent of next-of-kin of all clinically-affected mutation carriers, as approved by the local ethics committee. Clinical assessments and scanning were preformed between the years 1991 and 2005. All mutation carriers underwent comprehensive clinical and neuropsychological assessments including the MMSE (Folstein *et al.*, 1975) (see Appendix Two), and volumetric MRI scans at each visit (41 visits: 3-8 per subject). Each control subject had two MRI scans (except two subjects who had four scans each) adding up to 54 scans in total. The clinical status of each mutation carrier was classified at each time-point as: 1) FAD affected, if the subject fulfilled National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders (NINCDS-ADRDA) diagnostic criteria for probable AD (McKhann *et al.*, 1984) (see Appendix One) 2) MCI, if the subject fulfilled the MCI criteria (Petersen *et al.*, 1999), and 3) presymptomatic, if subjects fell short of both NINCDS-ADRDA and MCI criteria. All mutation carriers eventually became affected and so have a known date of onset, except one who has remained presymptomatic. In this case, the age of onset was estimated from the mean age of onset from other affected members in the subject's pedigree.

6.3.2.2.MRI Acquisition

In the period 1991-2000, imaging was performed on a 1.5T GE Signa scanner (General Electric Medical Systems, Waukesha, WI, USA) according to the standard protocol (see Appendix Three). In the period 2001-2005, volumetric coronal T_1-weighted MR imaging was performed on a different 1.5T GE Signa scanner running software version 5.8 (General Electric Medical Systems, Waukesha, WI, USA) using the following protocol: a inversion recovery prepared fast spoiled gradient-echo technique (256×256 matrix, FOV 24×18 cm, TR/TE/TI/NEX/FA = 14 ms/5.4 ms/650ms/1/15° yielding 124 contiguous 1.5-mm-thick slices).

6.3.2.3.Image processing

Volumetric image analysis was performed blind to subject details using the MIDAS software. Brain regions were extracted from surrounding tissue using a semi-automated technique and manually edited where necessary to obtain a whole brain volume (see Appendix Four). TIV was calculated in order to adjust for differences in head size between individuals as previously described (Whitwell *et al.*, 2001) (see Appendix Four).

Change in whole brain volume was calculated using the BSI (Freeborough and Fox, 1997). The BSI was performed first by registering each repeat scan to its baseline pair using an affine (12 dof) algorithm, and then calculating the volume change directly over the entire 3D brain-CSF interface. Annualised rates of whole-brain atrophy were calculated as a percentage of the baseline-adjusted brain volume for each scanning interval. Atrophy rates that span the change in MRI scanner were ignored, as the BSI may not be reliable, due to differences in image quality between scanners.

For hippocampal volume measurements, all scans were first registered to a standard brain template (MNI 305) (Mazziotta *et al.*, 1995) using six dof algorithm to reduce any variability in landmarks used in delineating the hippocampus. Repeat scans were then registered to corresponding baseline scans using a nine dof algorithm. Each hippocampus was manually traced using multiple views to include the cornu ammonis, gyrus dentatus and subiculum (see Appendix Four). Left, right and total (left+right) hippocampal volumes were calculated and adjusted for TIV as previously described. Changes in adjusted total hippocampal volumes were calculated by subtracting repeat from the baseline hippocampal volumes for each scan pair. Annualised rates of

hippocampal atrophy were calculated as a percentage of the baseline hippocampal volume for each scan pair.

6.3.3. Statistical analysis

Mean adjusted whole brain and total hippocampal volumes were calculated for each subject category (controls, presymptomatic, MCI and AD). The presymptomatic and AD stages were further subcategorised into earliest and latest stages, as most mutation carriers had more than one scan in each of these categories. Rates of atrophy across category transitions (last presymtomatic to MCI transition, MCI (or last presymptomatic, if no MCI stage identified) to earliest AD transition) and within categories (earliest presymptomatic to latest presymptomatic, earliest AD to latest AD) were also calculated. Imaging data of patients in each category were compared with controls using two-sample t-tests with allowance for unequal variances.

Separate hierarchical models were developed for adjusted whole brain and total hippocampal volumes using SAS Proc Mixed. The models related the repeated volume measures to time to clinical diagnosis of AD. Random intercepts and slope terms, where needed, were fitted to allow for the repeated within-subject measurements.

In order to compare volumes and atrophy rates between controls and mutation carriers, the date of a control's last scan was assumed to correspond to date of clinical diagnosis in mutation carriers. A constant atrophy rate for controls was assumed.

Fixed quadratic terms in time were allowed to assess evidence of acceleration in atrophy in the mutation carrier group. Where there was evidence it was needed and estimation was possible, separate variance parameters for the control and AD groups were fitted. Log-transformation of volume measurements was performed to model percentage changes. Since some patients had scans on two different scanners, a fixed scanner effect covariate was added. To account for the small sample size, adjusted degrees of freedom were used for the confidence intervals of fixed effects (Kenward and Roger, 1997).

A separate hierarchical model was fitted to the BSI-derived measures of brain volume changes between consecutive scans, allowing for the additional within-subject correlation structure of direct measures of change (Frost *et al.*, 2004). The outcome

variable was the logarithm of the ratio of volume at the second scan to the volume at the first scan, with the volume at the second calculated using the volume at the first and the BSI change.

6.3.4. Results

Table 6.3 shows that adjusted total hippocampal volumes of mutation carriers were smaller than those of age- and gender-matched controls at each stage, but the differences were not statistically significant until the MCI stage. There were statistically significant differences in adjusted whole brain volumes only once mutation carriers fulfilled the clinical criteria for AD. However, there was statistically significant evidence that rates of hippocampal atrophy and whole brain atrophy were already higher in the presymptomatic and subsequent MCI and AD stages for mutation carriers compared with controls (Table 6.4). As mutation carriers moved through the clinical stages of AD, both total hippocampal and whole brain mean atrophy rates correspondingly increased.

In the control group, there was statistically significant hippocampal asymmetry with the right hippocampus (mean±SD) larger than the left (2981±230 vs. 2815±177, $p<0.0001$). Although the mean right hippocampal volume was consistently larger in value than the mean left hippocampal volume in mutation carriers at every clinical stage, the left/right differences were not statistically significant (paired t-tests, all $p>0.05$). There was no evidence that unaffected relatives differed from the other controls in either volume or atrophy rates, for either total hippocampus or whole brain (all $p>0.2$).

Table 6.3 Cross-sectional analysis of the data according to clinical status.

	Controls, n=25	Presymptomatic, n=8		MCI, n=6	AD, n=7	
Subcategory*	Mean	Earliest	latest	mean	earliest	latest
Age (years), mean (SD)	46.5 (10)	43.4 (7.1)	44.7 (6.5)	49.4 (6.6)	49.8 (7.6)	50.8 (7.9)
Males, n (%)	9 (36)	3 (38)	3 (38)	3 (50)	3 (43)	3 (43)
Years from AD diagnosis, mean (SD)	NA	-4.6 (2.1)	-3.3 (2.1)	-1.8 (1.7)	0.5 (0.8)	1.5 (1.5)
MMSE, mean (SD)	NA	28.9 (1.1)	28.0 (1.7)	27.0 (2.5)	23.9 (4.0)	18.9 (7.8)
Adjusted brain volume(TR) (ml), mean (SD)	1212.6 (42.3)	1214.7 (49.6) p=0.92	1205.4 (48.6) p=0.71	1162.9 (71.6) p=0.15	1129.5 (54.0) p=0.005	1103.5 (74.0) p=0.007
Adjusted total hippocampal volume (mm3), mean (SD)	5797.1 (378.7)	5502.9 (710.7) p=0.29	5367.7 (748.8) p=0.16	4989.5 (780.3) p=0.05	4624.9 (741.9) p=0.005	4366.0 (925.2) p=0.006

P-values are for two-sample t-tests comparing patients in each category with controls.
** Earliest refers to the earliest scans/clinical assessment available in that stage; latest refers to the latest scans/clinical assessment available in that stage; mean refers to the mean values available from scans/clinical assessments available at that stage.*

Table 6.4 Longitudinal analysis of the data according to transition stage.

Transition*	Control-control N=25	presymptomatic-presymptomatic N=6	Presymptomatic-MCI* N=5	Presymptomatic /MCI-AD* N=6	AD-AD N=5
Age, (years), mean (SD)	46.5 (10.2)	42.9 (7.6)	49.0 (7.2)	48.0 (7.8)	52.1 (7.8)
Mean years from diagnosis of AD, mean (SD)	NA	-3.9 (1.6)	-2.6 (2.0)	-0.2 (0.5)	1.43 (1.03)
Males, n (%)	9 (36)	3 (50)	2 (40)	3 (50)	2 (40)
Mean MMSE, mean (SD)	NA	28.3 (1.3)	27.0 (1.4)	26.0 (2.1)	19.5 (3.9)
Scan interval (years), mean (SD)	1.5 (0.8)	1.8 (0.9)	1.5 (0.6)	1.5 (1.0)	1.3 (0.7)
Rate of MMSE decline (points/year), mean (SD)	NA	0.7 (1.1)	0.7 (2.2)	0.8 (0.9)	4.1 (4.3)
Brain atrophy rate (%/year), mean (SD)	0.00 (0.57)	0.51 (0.32) p=0.01	0.52 (0.51) p=0.08	1.82 (0.87) p=0.003	1.79 (0.74) p=0.004
Total hippocampal atrophy rate (%/year), mean (SD)	0.31 (1.25)	1.74 (1.12) p=0.03	3.02 (1.85) p=0.03	3.44 (1.25) p=0.0007	6.52 (2.31) p=0.003

P-values are for two-sample t-tests comparing patients in each category with controls.

** Comparison was between the latest scan/clinical assessment available in earlier stage and the earliest scan/clinical assessment of the later stage.*

Figure 6.3 shows the adjusted total hippocampal volume measurements for mutation carriers and controls. The hierarchical model showed strong evidence of hippocampal atrophy ($p<0.0001$) and acceleration ($p<0.0013$) in the mutation carrier group. The estimated mean difference in adjusted total hippocampal volume between mutation carriers and controls therefore increased with time, with the difference becoming statistically significant at the 5% significance level just over three years before fulfilling the clinical criteria for AD (Figure 6.4). By that time, the mean total adjusted hippocampal volume of mutation carriers was estimated to be 90.16% (95% CI 81.72%, 99.47%) that of controls.

The model estimated the annualised total hippocampal atrophy rate in controls to be 0.23% (95% CI -0.19%, 0.66%). The estimated difference in the mean hippocampal atrophy rate between mutation carriers and control groups became statistically significant (5% level) 5.5 years prior to the time of clinical conversion to AD, with a difference of 1.55% (95% CI 0.04%, 3.04%) (see Figure 6.5). The mean hippocampal atrophy rate difference was estimated to increase by 0.40% (95% CI 0.17%, 0.63%) per year.

Figure 6.6 shows the adjusted whole brain volumes for mutation carriers and controls. The model for adjusted whole brain volumes showed strong evidence of increasing atrophy in the mutation carrier group ($p<0.0001$). The estimated mean difference in the adjusted whole brain volume between mutation carriers and controls thus increased with time, with the difference becoming statistically significant one year before clinical diagnosis (see Figure 6.7), by which time, the mean whole brain volume of mutation carriers was estimated to be 96.36% (95% CI 92.95%, 99.91%) that of controls. There was evidence ($p<0.0001$) of a scanner effect in this model with a mean difference of 4.77% (95% CI 3.21%, 6.35%) in brain volume measurement between the two scanners, with the former providing larger volumes.

From the hierarchical model for BSI-derived whole brain changes, the annualised rate of atrophy in controls was estimated to be 0.01% (95% CI -0.14, 0.16%). There was strong evidence of whole brain atrophy in the mutation carriers ($p<0.001$), and also of acceleration in the whole brain atrophy rate ($p<0.001$). The estimated difference in mean whole brain atrophy rates between mutation carriers and controls became statistically significant around 3.5 years prior to AD diagnosis, with a mean difference

(3.5 years prior to conversion) of 0.51% (95% CI 0.12%, 0.89%) (see Figure 6.8). The mean difference increased by 0.26% (95% CI 0.16%, 0.37%) per year.

There was no evidence that either total hippocampal or whole brain atrophy rates differed according to gender at each of the clinical stages within the mutation carrier group (unpaired t-tests; all p>0.05), or from the hierarchical models for total hippocampal volume and whole brain volume (both p>0.2). There was statistically significant evidence from the model for BSI-derived whole brain changes that the atrophy rate was greater male mutation carriers than in women (p=0.03), although with such small numbers this may well be due to confounding.

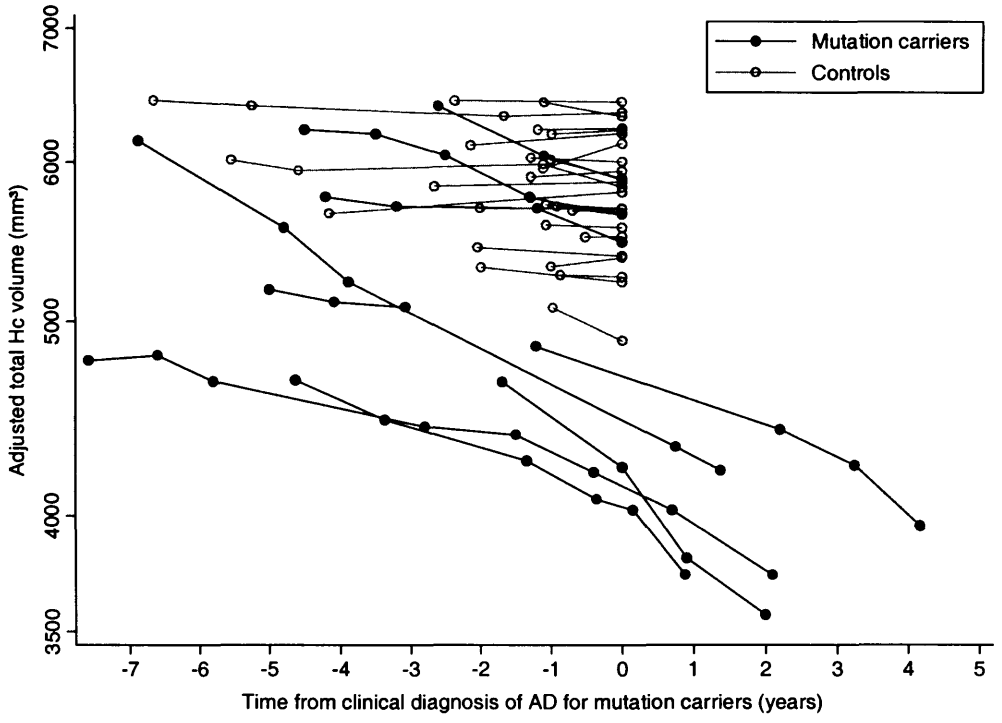


Figure 6.3 Adjusted total hippocampal volume measurements of mutation carriers (relative to time of clinical diagnosis of AD) and controls (relative to the date of their last scan).
The y-axis scale is logarithmic. AD= Alzheimer's disease.

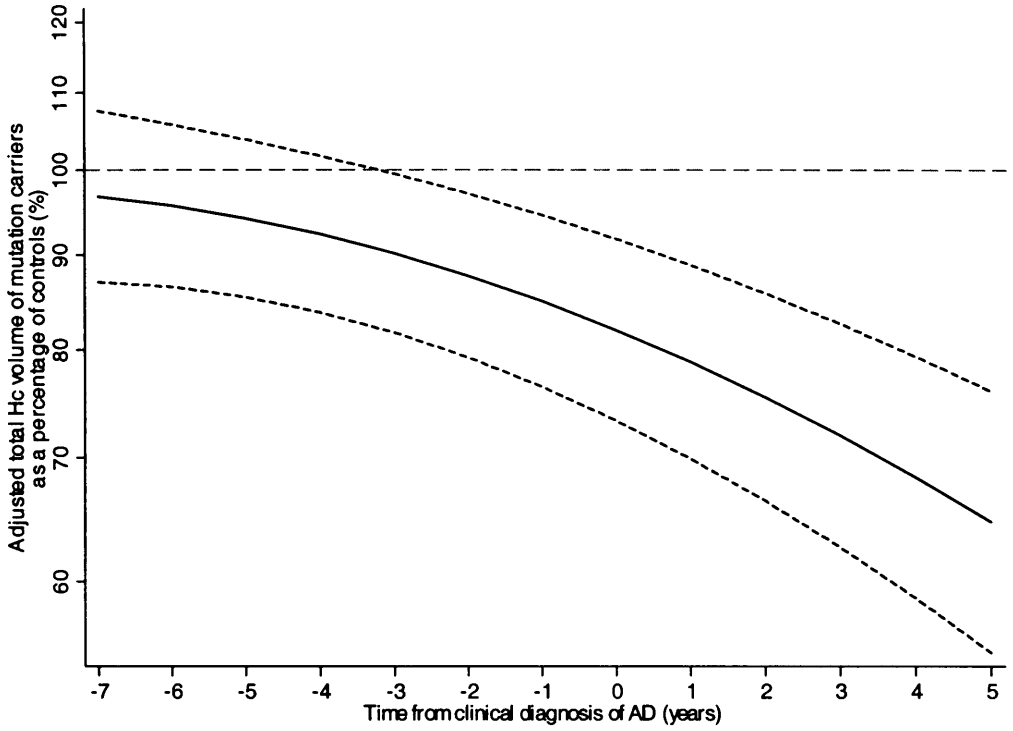


Figure 6.4 Model estimated mean adjusted total hippocampal volume of mutation carriers as a percentage of mean adjusted total hippocampal volume of controls. Short dashed lines indicate approximate 95% confidence interval limits for the mean difference. The y-axis scale is logarithmic. Hc = hippocampus.

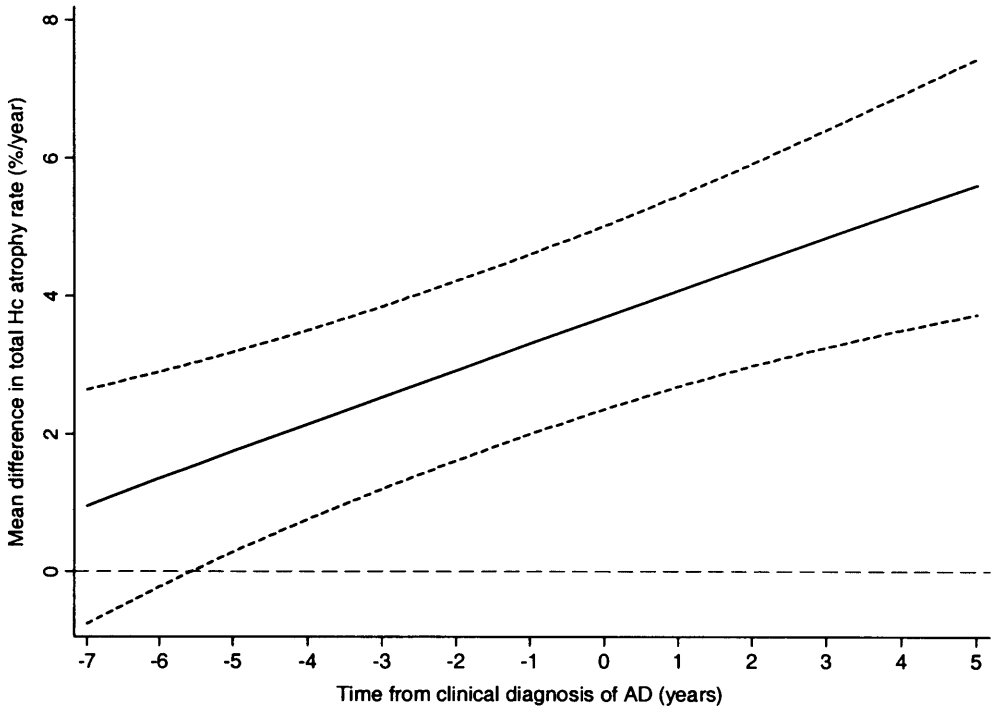


Figure 6.5 Model estimated mean difference in total hippocampal rate between mutation carriers and controls.
Short dashed lines indicate approximate 95% CIs for the mean difference.
Hc = hippocampus

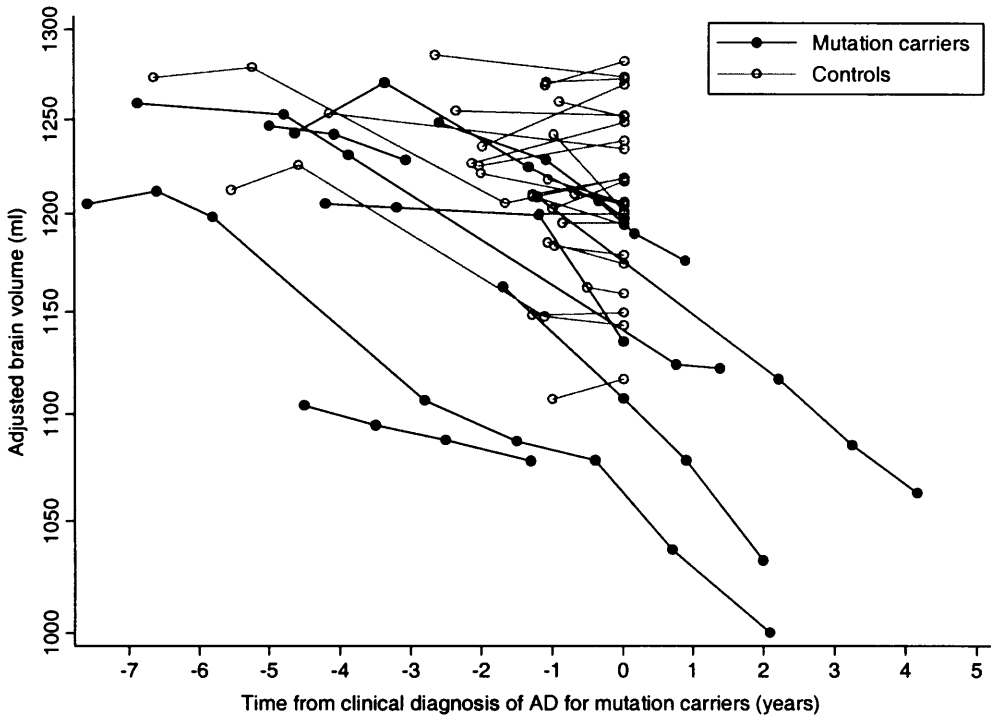


Figure 6.6 Adjusted whole-brain volume measurements of mutation carriers (relative to time of clinical diagnosis of AD) and controls (relative to the date of their last scan). The y-axis scale is logarithmic.

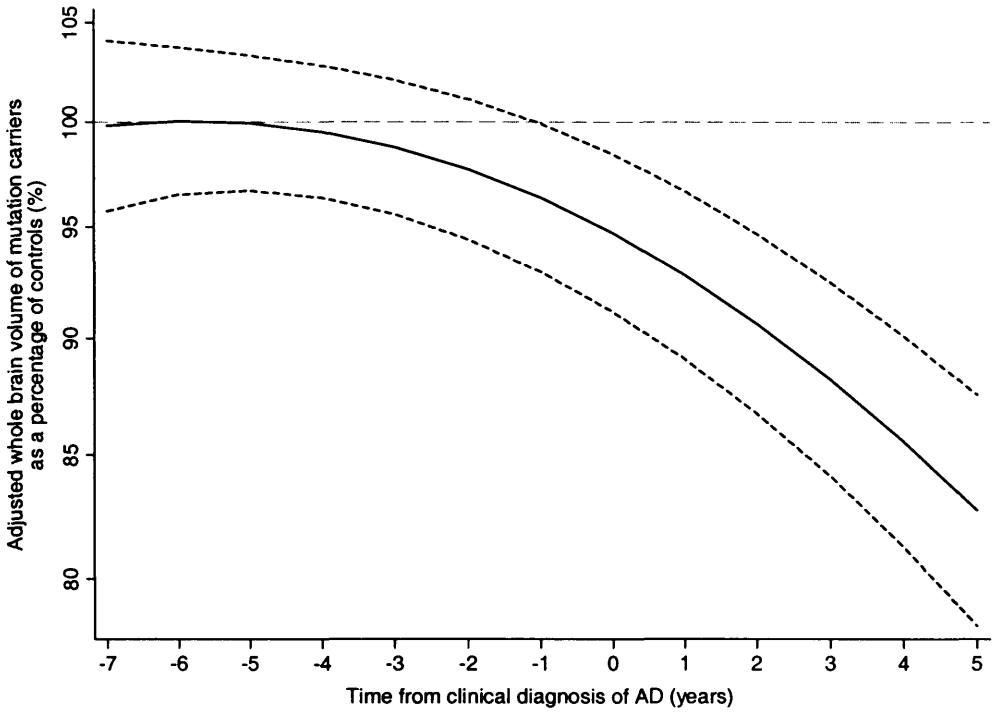


Figure 6.7 Model estimated mean adjusted whole-brain volume of mutation carriers as a percentage of mean adjusted whole-brain volume of controls.
Short dashed lines indicated approximate 95% CIs for the mean difference. The y-axis scale is logarithmic.

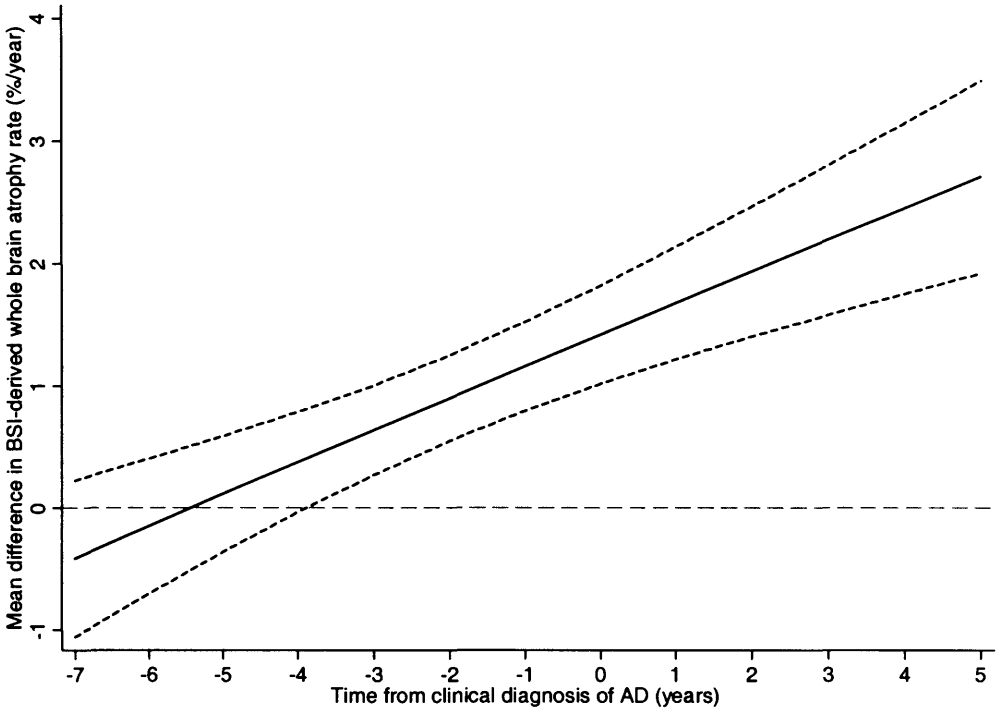


Figure 6.8 Model estimated mean difference in BSI-derived whole brain atrophy rates between mutation carriers and controls.
Short dashed lines indicate approximate 95% confidence interval limits for the mean difference.

6.3.5. *Discussion*

This study demonstrates evidence of atrophy of whole brain and hippocampus as seen on serial MRI, which predates the clinical diagnosis of AD. The study has shown that by the time AD was diagnosed clinically, the estimated mean hippocampal volume in patients was 18.1% smaller than in controls and mean whole brain volumes were 5.4% smaller. Longitudinal measures of hippocampal and whole brain atrophy were more sensitive than cross-sectional volume measures at detecting group differences, with the former detecting change 2-3 years earlier than the latter. This is somewhat unsurprising since cross-sectional data are more susceptible to inter-subject variability than longitudinal measures. By using more mutation carriers and assessment/scanning data points than previously reported, this study was able to 1) demonstrate statistical evidence of acceleration in the atrophy rates of whole brain and hippocampus in the years preceding and following the diagnosis of AD, and 2) provide more precise estimates of volumes and rates of atrophy at different clinical stages of the disease than previously published (Schott *et al.*, 2003).

In this study, there was no evidence of progressive atrophy in the control subjects, which probably reflects their young age (mean=46 years). Previous work has shown evidence of whole brain and hippocampal atrophy as a result of normal ageing with acceleration above the age of 70 years (Scahill *et al.*, 2003).

Although the hippocampal region is not the only region of interest in the evolution of AD neuropathology, the hippocampus was chosen as an example of a region that has been shown to be affected early in the disease process, and it was monitored as to how it changed in relation to the clinical stage. Understanding the involvement of other brain regions remains an important issue to be addressed in future studies. This would require either manual segmentation of other brain regions (such as the entorhinal cortex) or evaluating the pattern of atrophy without prior assumptions of where the pathology starts (for example, using voxel based morphometry).

There are a few limitations to this study. Generalisation of our findings based on FAD cases to sporadic AD is limited by the potential biological differences in the two forms of the disease, with the former having definite genetic risk factors and significantly younger age at disease onset than the sporadic form. Nonetheless our findings accord

with previous studies in sporadic AD and MCI which have shown atrophy of medial temporal lobe structures and whole brain prior to clinical diagnosis (Fox and Schott, 2004; Jack *et al.*, 2000; Jack *et al.*, 2004; Kaye *et al.*, 2005; Rusinek *et al.*, 2003; Stoub *et al.*, 2006). One advantage of the study of individuals at risk for FAD is that it is feasible to study asymptomatic subjects longitudinally for many years prior to clinical onset. Another strength of this cohort is that their young age means they have relatively little age-related co-morbidity such as vascular disease that may confound assessment of the earliest effects of AD pathology in older sporadic AD.

One potential source of bias in our study is the exclusion of five other scans among mutation carriers due to sub-optimal quality, which could have provided further data points. Another source is that control subject had less scans per subject than mutation carriers. As a result, a constant atrophy rate in controls had to be assumed and these results extrapolated under the assumption they are constant. Although having fewer scans results in less precise estimates, attempts were made to improve the estimates by including more controls in comparison to mutation carriers (25 vs. 9).

It is acknowledged that although this is the largest analysis of serial imaging in FAD reported (n=9 having a total of 41 scans/assessments), the sample size is still relatively small. In addition, although imaging resolution was consistent, scanning was done on two units with different sequence parameters. This is inevitable due to scanner and sequence upgrades over the 14 years of assessment. An attempt was made to control for this by avoiding BSI measurements across scanners and by including scanner-unit as a covariate in our statistical models. Finally, hierarchical models that allowed volumes to vary as a quadratic function of time were fitted, which is clearly a simplification of the truth. Estimates are conditional on this parameterisation, and therefore should be viewed with appropriate caution. However, this is more realistic than a model assuming constant rates of atrophy (Schott *et al.*, 2003). With larger datasets it will be possible to fit more sophisticated patterns of atrophy.

In conclusion, this study provides evidence of acceleration of brain atrophy with disproportionate hippocampal involvement preceding the clinical diagnosis of AD in mutation carriers.

6.4. Chapter conclusions

This chapter shows the use of combining cross-sectional and longitudinal data in investigating both subtle questions such as asymmetry and to establish when changes can be detected in hippocampal volumes prior to the development of symptoms.

7. AUTOMATED SUBDIVISION OF MANUAL SEGMENTATIONS

7.1. Chapter introduction

Subdivision of structures

As described earlier in Chapters 4 and 5 (pages 98 and 134), there is evidence to suggest that pathological processes such as AD affect the cingulate at an early stage of the disease. Moreover, evidence from a number of imaging modalities and pathological studies have suggested that some sections may be more affected than others; PET imaging has revealed hypometabolism principally in the posterior sections, and MRS has shown metabolite changes again in the posterior, and volumetric MR imaging has revealed differential patterns of volume of subsections along the length of the cingulate (Callen *et al.*, 2001; Killiany *et al.*, 2000) (see results presented in Chapter 4, page 108 and Chapter 5 (page 148)). Measuring atrophy in sections of the cingulate may therefore be a more useful and sensitive marker of the disease, particularly in the early stages. In addition, understanding which section is affected may mean reduction in the amount of labour-intense delineation of regions in the future as one section may require segmentation.

Manual segmentation

It has been shown that subdivision of the cingulate region into smaller ROIs reveals patterns of atrophy which differ between patient groups (see Chapter 4 (page 109) and Chapter 5 (page 148)). However, this requires an initial segmentation of the cingulate gyrus which takes approximately fifty minutes followed by manual subdivision taking a further ten minutes. Manual subdivision requires application of subdivision protocols and much like any other protocol, the interpretation of this may introduce operator errors. Furthermore, the cingulate is a complex structure with convoluted folds, together with a large degree of inter-individual variation in gyral and sulcal patterns within and proximal to the gyrus. As described in detail in Chapter 4 (page 98), this makes consistent labelling and subdivision difficult as the manual subdivision is dependent on landmarks which are not proximal or intrinsic to the structure; the anterior and posterior commisures are located at some distance from the gyrus and as a result the application of protocols dependent on these structures could potentially introduce bias, and/or variability or inaccuracy in these results. Therefore subdivision of the cingulate according to approximate tissue types as described in Chapter 4 and Chapter 5 (see pages 106 and 141 respectively) may be unnecessarily effortful.

As discussed in Chapter 4 and Chapter 5, subdivision of regions may be appropriate in many cases; however scale must be considered. The care taken per voxel when segmenting small structures as a proportion of the whole structure is much greater than the care taken with larger structures. Therefore, this form of subdivision may not be appropriate in all cases and instead, further manual demarcation protocols may be developed for the smaller structures of interest for example, the retrosplenial cortex.

Other methods used for subdivision

Other structures which may be considered to be similar to the cingulate in structure have been subdivided automatically. The corpus callosum is a white matter structure which runs along the inferior of cingulate gyrus which joins the two cerebral hemispheres. This structure has been studied in AD by segmenting it on the mid-sagittal slice and subdividing into a number of sections using a template mask see Figure 7.1 (Hampel *et al.*, 1998). Problems with accuracy may arise when using this type of method as it is theoretically not robust to differences in the relative anatomy (see Figure 7.2) which are particularly extensive in the cingulate (Ono *et al.*, 1990) (see Figure 7.3). Instead, for the cingulate, a method which models the curve of the structure and divides according to its relative length may be more appropriate and potentially less susceptible to the diameter of the structure affecting the borders of the subdivisions.

Bezier method application to cingulate gyrus

In this chapter a method of subdivision of the cingulate into distinct sections i.e., RAC, CAC and PC using a Bezier curve method of automated subdivision is described and assessed. This chapter aims to describe the application of this technique and the potential advantages in using this for structures such as the cingulate gyrus.

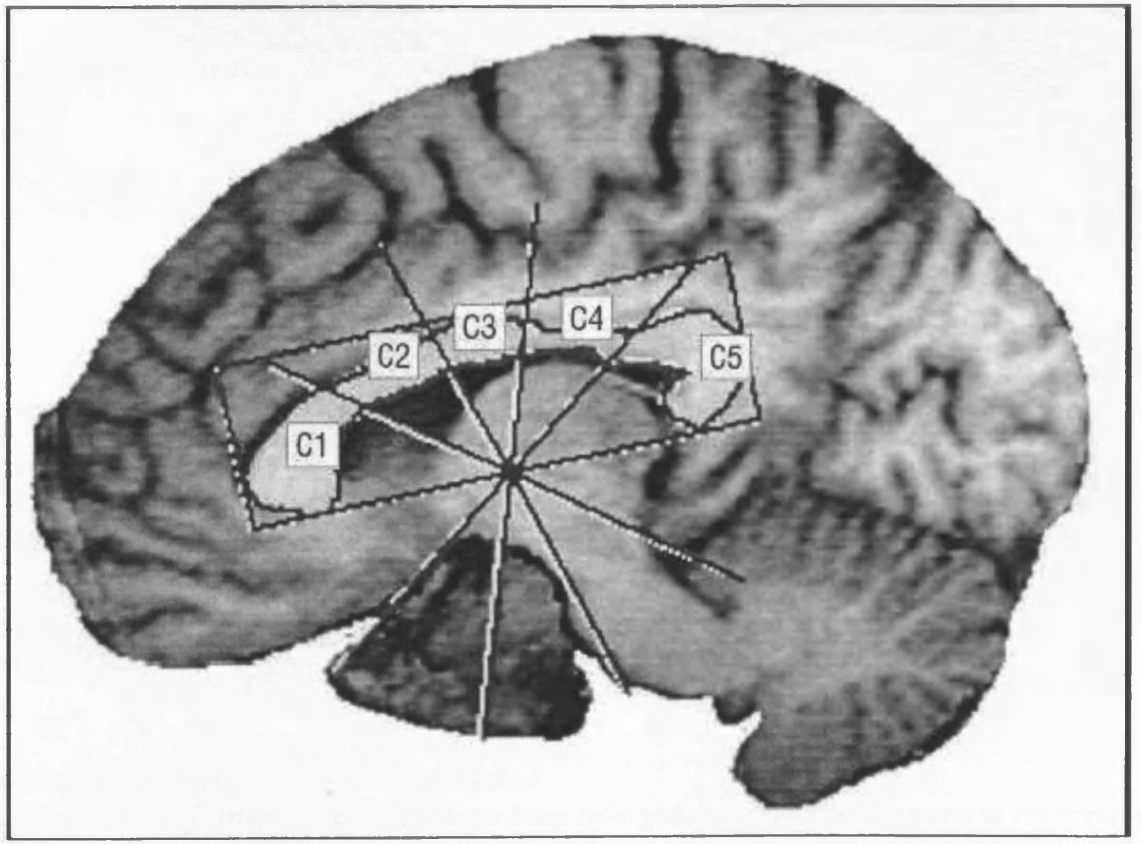


Figure 7.1 Subdivision of the corpus callosum using Hampel's box method.
Measurement of the corpus callosum subregions is performed in the midsagittal slice (c1: rostrum and genu; c2–c4: truncus; c5: isthmus and splenium).

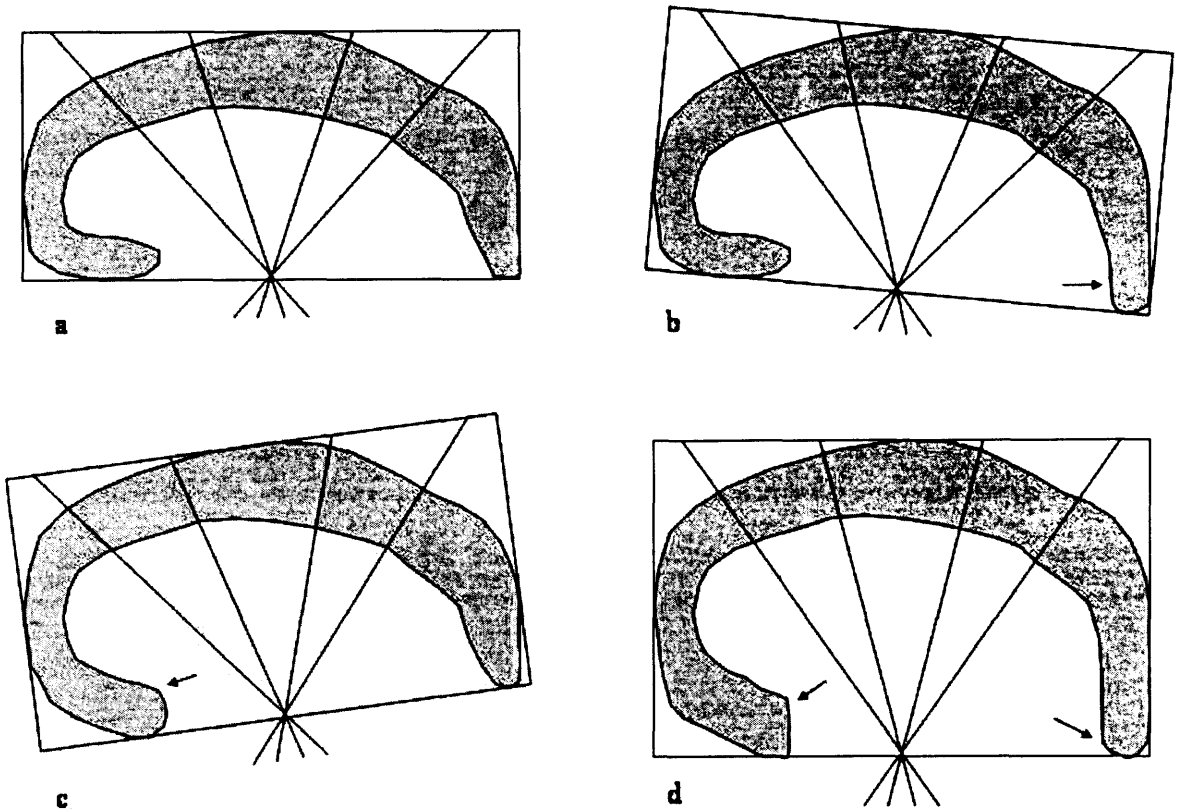


Figure 7.2 A demonstration of the problem of Hampel's box method.
Structures 'b-d' are edited variations of structure 'a' which have been enlarged and extended inferiorly in regions marked by arrows. By changing the diameter of the structure at the inferior border (arrows), the template box rotates and the segmentation of all subdivisions is influenced.

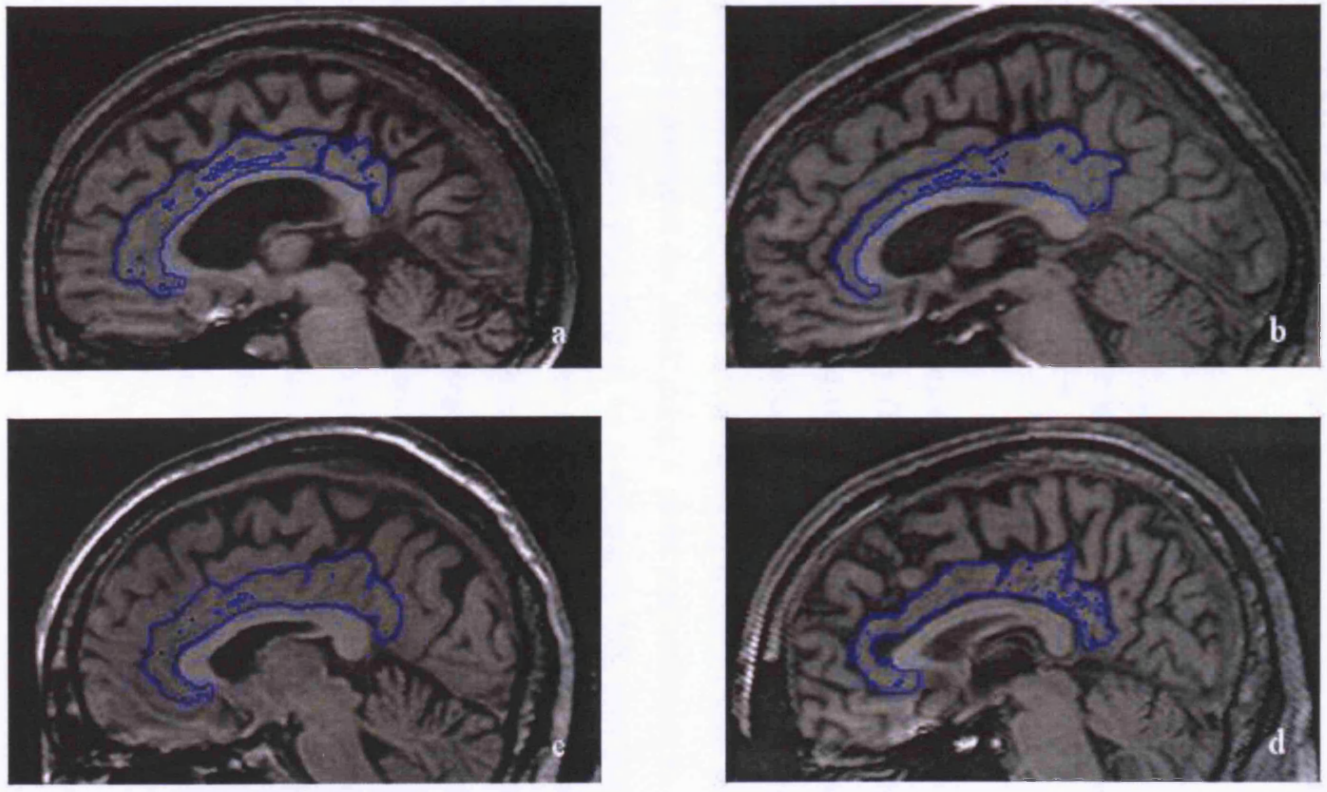


Figure 7.3 Segmentations of the cingulate gyrus.
This shows variety of different types of topology in the cingulate region in four different subjects (a-d).

7.2. *Methods*

7.2.1. *Subjects and Imaging*

Nineteen AD and eleven control subjects with mean ages approximately 56 years were included in this study. For further details of these subject groups see Chapter 5 (see page 135 and Table 5.1). Each subject had two scans taken at over one year inter-scan interval.

7.2.2. *Cingulate segmentation*

The manually-delineated cingulate ROIs which were used for subdivision were described in Chapter 5 (page 134).

7.2.3. *Cingulate subdivision*

The Bezier method of subdivision by (Boyes *et al.*, 2004) was used. Briefly, the cingulate is segmented on T1-weighted volumetric MRI scans to define a binary mask that identifies the structure on the scan (see Figure 7.3). The method finds an approximation to the central axis of the cingulate by fitting a Bezier curve through the volume and then partitions the mask using a plane perpendicular to this axis. This method takes a cingulate ROI defined on a volumetric MR scan and fits (in a least squares sense) a cubic Bezier curve to this mask. This curve is used as a smooth approximation to the central axis of the cingulate, enabling the estimation of the length of the structure, and also the partitioning of the three dimensional mask into sections. The fitting process requires that start and end points of the curve are manually fixed. These were selected to be the lowest point of the posterior cingulate and in the middle of the anterior cingulate on the same sagittal slice (see Figure 7.4). The cingulate was then divided into thirds on the baseline and repeat ROIs. Rates of atrophy were calculated in the way described previously (see Chapter 5, page 141).

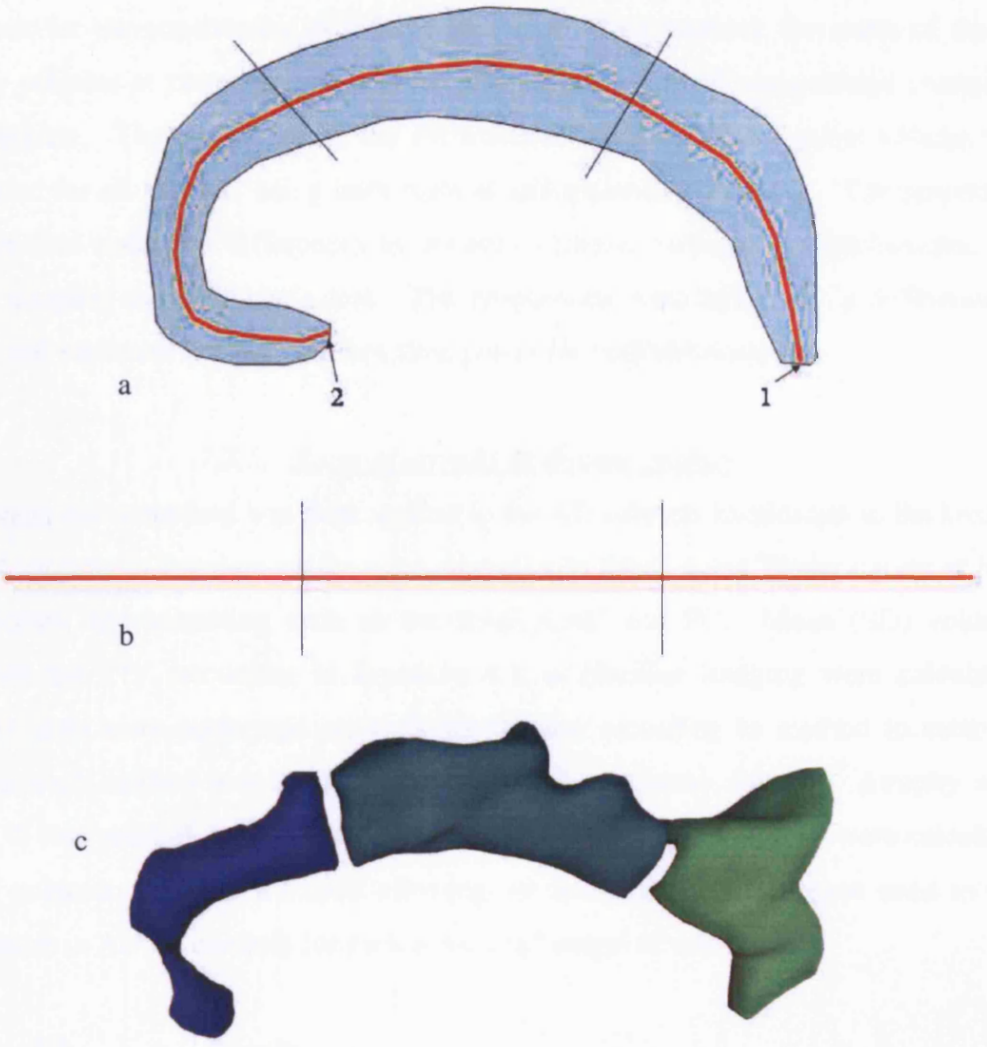


Figure 7.4 Schematic diagram showing landmarks 1 and 2 are manually selected by the use for Bezier-based cingulate subdivision.
'a' shows the curve of the cingulate is modelled by a Bezier curve. The user selects which region of the ROI to subdivide. The subdivision is based on the relative length of the curve 'b' which is selected by the user. The result 'c' is a 3D rendered example of a subdivided ROI.

7.2.4. *Testing for consistency*

As a test for the consistency of the Bezier curve fitting method, the scans of the 11 healthy subjects at two time points were used, assuming no disease-related change in the cingulate. The proportion of the PC volume to the whole cingulate volume was calculated for all subjects using both manual and automated methods. The proportion measure was tested for differences in variance between methods at both baseline and repeat imaging using Pitman's test. The proportions were tested for a difference in means and variances between the two time points for both methods.

7.2.5. *Rates of atrophy in disease groups*

The Bezier curve method was then applied to the AD subjects in addition to the healthy control subjects. Cingulate masks were cleaved into thirds using Bezier curves at both time points, approximating each as the RAC, CAC and PC. Mean (SD) volumes corrected for TIV, according to Equation 4.1, at baseline imaging were calculated. Paired t tests were performed on each subdivision according to method to establish whether each method was sampling approximately the same volume. Atrophy rates (mean % loss/year) of the whole cingulate and the three divided regions were calculated for all subjects. Student's t tests allowing for unequal variances were used to test differences in AD vs. controls for each automated cingulate subdivision.

7.3. *Results*

7.3.1. *Consistency*

The initial test comparing the Bezier curve cleaving method with manual segmentation of the PC showed that the difference in the variance in PC as a proportion of cingulate in the control group was reduced using the automated method at both time points (baseline ($p=0.009$) and repeat imaging ($p = 0.013$)). There was no evidence that this proportion was less variable over time ($p = 0.53$) using the automated measurement.

7.3.2. *AD and controls across the subdivisions*

Volumes of the cingulate on baseline imaging according to manual and automated methods are presented in Table 7.1 with associated p values. Mean volumes were significantly different in each subdivision according to method except for the posterior cingulate in controls. At baseline, only the PC was significantly different in volume between the AD and control group ($p=0.0051$ for manual methods and $p=0.00072$ for automated methods). The rates of atrophy of the cingulate and its sub-regions were

different in means between AD and control subjects (Table 7.2 and Figure 7.5). Within the AD group the PC had a higher rate of atrophy than the whole cingulate, although it was not significant (p = 0.078).

Table 7.1 Mean (SD) volumes of the cingulate gyrus in AD and controls according to method at baseline imaging.

	Controls			**AD**		
	Manual	Automated	p value	Manual	Automated	p value
RAC	7733 (1426)	4568 (825)	<0.0001	7168 (1400)	4543 (1080)	<0.0001
CAC	3861 (1043)	5910 (1298)	0.0006	3409 (635)	5135 (812)	<0.0001
PC	5594 (1166)	6560 (1297)	0.15	4324 (1067)	5140 (1044)	<0.0001

Table 7.2 Mean and (SD) rates of atrophy by partitioning cingulate automatically using Bezier curves.

	Cingulate	**RAC**	**CAC**	**PC**
Controls	-0.3 (1.2)	0.5 (6.5)	-0.5 (3.0)	-2.2 (3.8)
AD	5.9 (3.5)	4.2 (5.4)	5.5 (4.4)	8.1 (5.4)
p-values	< 0.0001	0.12	0.0001	< 0.0001

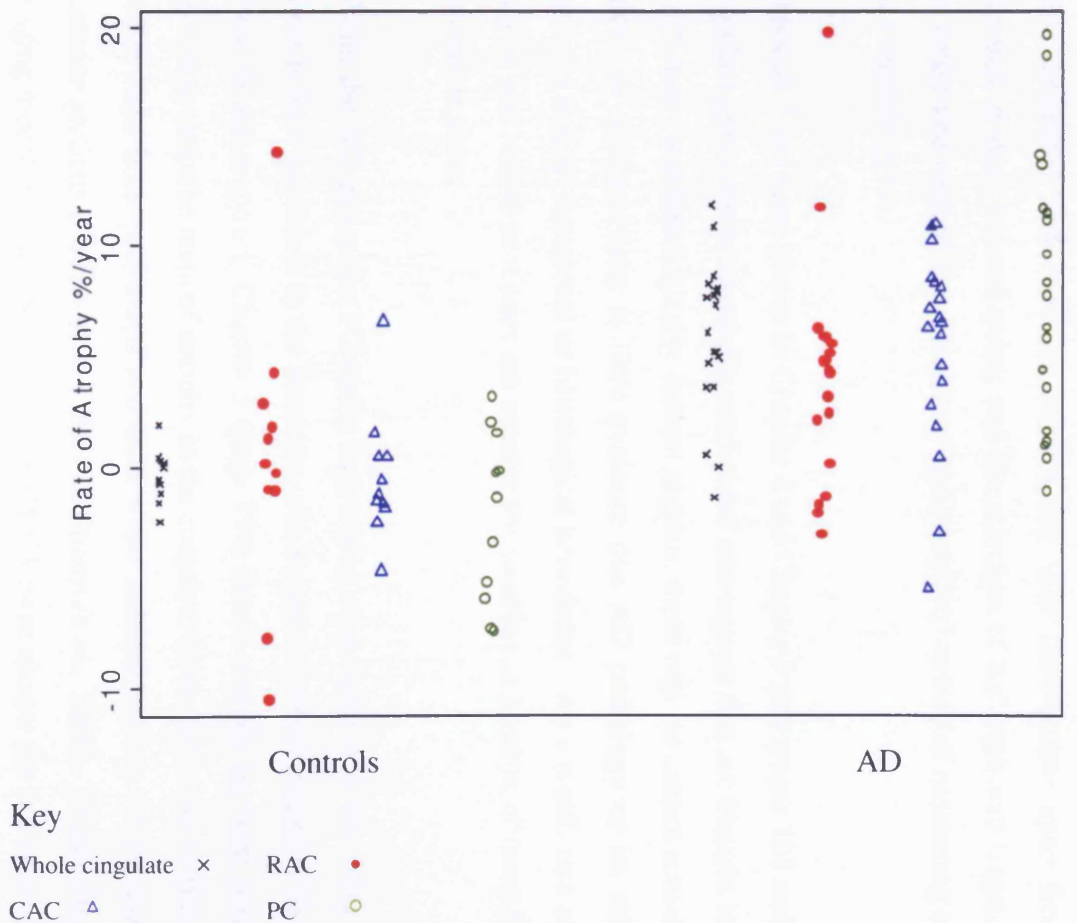


Figure 7.5 Rates of atrophy in the whole cingulate and its three major subdivisions.

7.4. Discussion

This study has shown that subdivision of the cingulate can be automated and the Bezier method is provides more consistent proportions of the cingulate than manual subdivisions. However, as the mean volumes of each subregion differed according to method this may be the result of differing regions being sampled by each method. For example, the cleaved PC region was larger and therefore included areas which were generally less variable. Another explanation may be due to the fact that the method is fully automated and does not require any user intervention apart from the initial selection of start and end points and the selection of the fractional length. Altogether, the automated method is a consistent and reproducible way of measuring subsections of the cingulate gyrus.

Although it has been shown in Chapter 4 and Chapter 5 (see pages 106 and 141) that the cingulate gyrus can be manually subdivided into regions that are thought to approximate functionally and histologically distinct regions, these regions cannot actually be seen on MRI. In addition there is little evidence that AD pathology or its atrophic effects perfectly respect functional or histological boundaries. As a result, this sort of method may be less biased as it does not assume the location of borders of these histological or function regions.

Within the AD group the PC had a higher rate of atrophy than the other subdivisions. There is little published in the literature with respect to atrophy rates of the subdivisions of the cingulate gyrus. Chapter 5 (page 134) details a study on these patients which is the first to describe rates of atrophy in the cingulate however, cross-sectional volumetric studies have been performed and both show reduced volume of the cingulate in the posterior sections (Callen *et al.*, 2001; Killiany *et al.*, 2000). Studies involving other imaging modalities including PET and SPECT have shown the posterior cingulate to be affected in AD (Matsuda *et al.*, 2002; Minoshima *et al.*, 1997; Valla *et al.*, 2001).

Improvements to the method's application could be made by performing the analyses on a more consistent scan interval and protocol, which may reduce the amount of variation in the measurements. Fitting Bezier curves to the cingulate provides an accurate approximation to its central axis as the structure is long, reasonably thin and has a definite start and end point. Irregularities such as concavities and/or convexities are

smoothed out as the ratio of these to the overall length of the structure are relatively insignificant. Any large deformations in the structure would have an adverse affect on the final fit, however. This method may be unsuitable for other structures if the central axis is not as apparent as that of the cingulate. An ideal structure for subdivision using this method would be the corpus callosum as it has a well defined central axis, and is also smooth. Although it is assumed that this method of subdivision may be less biased to the anatomical differences in structure borders as Hampel's method, this study did not assess this and as a result this remains purely a theoretical advantage.

7.5. Chapter conclusion

Automation may allow more precise subdivisions of structures to be made from larger ROIs with little extra manual intervention. These methods may reveal different patterns of atrophy across structures.

8. SEMI-AUTOMATED METHODS OF MEASURING CHANGES IN THE HIPPOCAMPUS

8.1. Chapter Introduction

Background

The previous work (Chapters 5 and 6) describing segmentation of the hippocampus using the "gold standard" measure of atrophy requires the outlining of the ROI on each MRI "slice". This technique is time-consuming, and errors in labelling the hippocampus may reduce the sensitivity of this method to detect group differences. In cross-sectional analysis, the interpretation of differences from controls (e.g. for diagnosis) is limited by the large degree of inter-individual variation in brain morphology (see Chapter 4, page 97). Examination of volumetric change over time within individuals using serial scanning reduces problems of inter-individual variability and may provide better group discrimination (Fox *et al.*, 1996a) as well as providing a means to monitor disease progression. However, errors in manual delineation become very important when small differences are being measured; the hippocampus typically loses 3-6 % of its volume over a year in AD (Cardenas *et al.*, 2003; Jack *et al.*, 1998) which is the same order of magnitude as delineation errors in hippocampal measurement.

As a result there has been interest in developing methods for automated segmentation of the hippocampus, for example (Andreasen *et al.*, 1996; Csernansky *et al.*, 1998; Gosche *et al.*, 2001) (see Table 2.1). Fully automated methods to delineate baseline and follow up scans using a template often rely on the identification of a number of landmarks, which may also be time-consuming and still subject to errors. An alternative approach is to use a manually segmented region on a baseline scan to derive subsequent ROIs, which can be used to calculate rates of atrophy.

Chapter objectives

Until the automated techniques as described in Table 2.1 become fully validated, semi-automated approaches may be the most appropriate approach to measuring change in the hippocampus. This chapter investigates the use of linear registration of serial hippocampi followed by either application of the boundary shift integral (BSI), or a further non-linear registration step to calculate the deformation in hippocampal region, to measure change over the scanning interval.

8.2. *BSI-based methods of automating change*

8.2.1. *Introduction*

One semi-automated approach to measuring change in the brain over time is application of the BSI (Fox and Freeborough, 1997) following linear registration to accurately align brain structures. This method was designed for the calculation of change in both whole brain (Freeborough and Fox, 1997) and has also been applied to ventricular regions (Ezekiel *et al.*, 2004; Schott *et al.*, 2005). Both regions involve boundaries which exclusively involve CSF and brain. This section describes the first study which applies the BSI to smaller grey matter structures which are both complex in shape (see Appendix Figure 3) and have borders of mixed tissue types: grey matter, white matter and CSF.

In this study this technique was applied to the hippocampus to investigate its utility in measuring change in this smaller structure where pathology is known to be concentrated, especially at the early stages of the disease. Where the BSI was measuring change was assessed, assuming this to be predominantly the grey matter – CSF border of the hippocampus. This was compared this semi-automated measure of hippocampal change with conventional manual outlining of the hippocampus on serial MRI. These measures were applied to subjects with AD and normal controls in order to assess their diagnostic utility.

8.2.2. *Materials and Methods*

8.2.2.1.*Subjects*

A group of 32 AD subjects and 47 controls who each had two volumetric MRI scans were selected, and were matched for age, sex and scan interval. All patients fulfilled NINCDS-ADRDA criteria for probable AD (see Appendix One). In addition, three subjects had histological confirmation of AD and 12 were from families with familial AD and had genetic confirmation. Table 8.1 shows the subject demographics.

Table 8.1 Subject demographics.

	Control	AD
Number	47	32
M/F	25/22	14/18
Mean (SD) age in years	59.2 (14.1)	59.5 (11.3)
Mean (SD) MMSE	29 (1)	19 (5)
Mean (SD) scan interval in days	423 (176)	397 (134)

8.2.2.2.Image Acquisition

MR imaging was performed on a 1.5T GE Signa Unit (General Electric Medical Systems, Milwaukee, WI). Fifty-seven subjects (36 controls and 21 AD subjects) were scanned using the standard protocol, and 22 subjects (11 controls and 11 AD subjects) were scanned using the MIRIAD protocol (see Appendix 3).

8.2.2.3.Manual Hippocampal Segmentation

Scans were transferred to a Sun workstation (Sun Microsystems Inc., Mountain View, CA). The software package MIDAS was used for all manual segmentation (Freeborough *et al.*, 1997). The hippocampus was segmented according to a previously validated protocol (see Appendix 4). No time limits were imposed, and the average segmentation time per hippocampus was 45 minutes.

8.2.2.4.Hippocampal BSI

Figure 8.1 and Figure 8.2 summarise the series of steps which demonstrate how the BSI measure was derived. Serial brain images were first registered (spatially matched using nine degrees of freedom (dof)) applying whole brain to whole brain matching, details of the registration technique have been described previously (Woods *et al.*, 1998). The resulting rigidly registered-repeat scan B was then registered locally onto scan A using six dof, resulting in accurate spatial matching of the hippocampi. The region created by

subtracting an eroded (by one voxel) intersection region from a dilated (by one voxel) union region of the resulting two brain regions (brain region A and locally registered brain region B) was generated and a dilated (by one voxel) baseline hippocampal region was used to mask this to allow calculation of the BSI in this region of interest area, **E**. This gives a boundary region containing the intensity shift of the hippocampal border.

The BSI was calculated by evaluating a sum over the group of voxels in the region **E** of the two scans (see Figure 8.1). It was calculated as a volume, Δv, effectively the 3D integral of all the boundary shifts (see Figure 8.2 for a example of an idealised one dimensional boundary shift):

$$\Delta v = \frac{K}{I_1 - I_2} \sum_{x,y,z \in \mathbf{E}} \left(clip(I_{base}(x,y,z), I_1, I_2) - clip(I_{reg}(x,y,z), I_1, I_2) \right)$$

Equation 8.1

Where

$$clip(a, I_1, I_2) = \begin{cases} I_2 (a < I_2) \\ a (I_2 \le a \le I_1) \\ I_1 (a > I_1) \end{cases}$$

Equation 8.2

and K is the voxel volume, I_{base}, I_{reg} are the normalised baseline and registered repeat intensities, I_2, I_1 define the bounds of the clipping function. Using I_2 and I_1 a window size (I_2 - I_1), I_w, and window centre ((I_2 + I_1)/2), I_c, was calculated which are user defined input parameters to the BSI calculation as detailed previously (Freeborough and Fox, 1997).

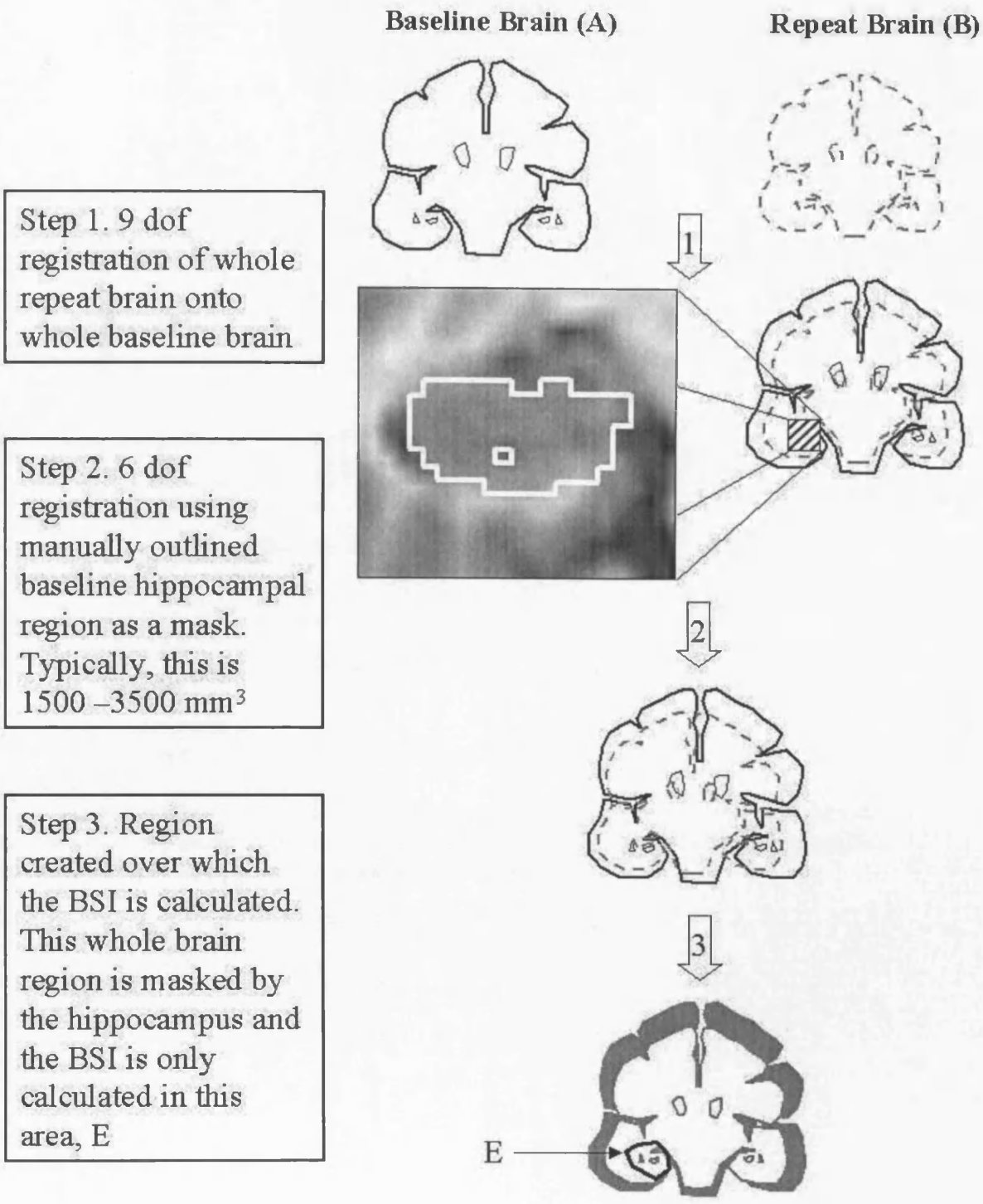


Figure 8.1 Summary of steps necessary to generate HBSI.
This demonstrates the registration steps leading to HBSI generation, involving two registration steps, first nine degrees of freedom (dof) to match whole brains and then six dof to match hippocampi. Following this, the region that contains the hippocampal boundary shift is generated and the HBSI is calculated in this region only.

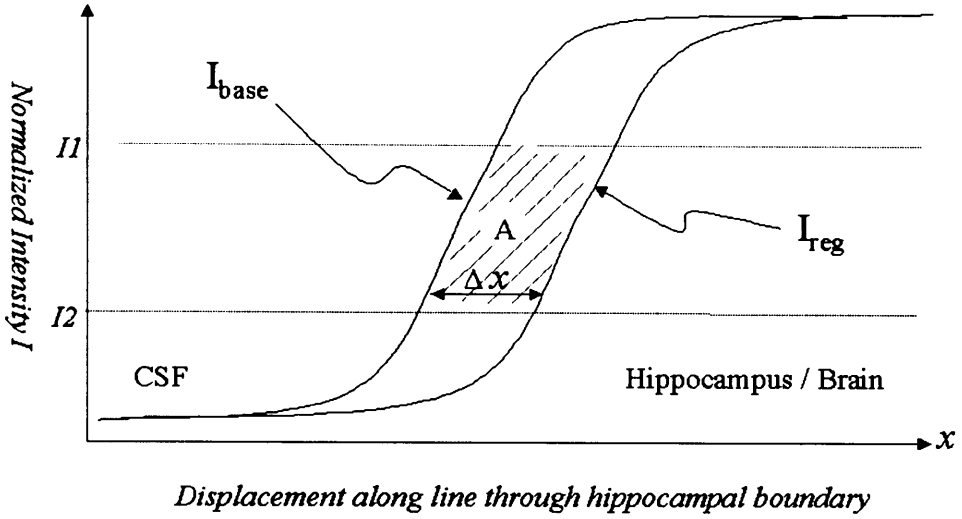


Figure 8.2. Example of an idealised one dimensional boundary shift in the hippocampus.
Area A is divided by the span of the intensity window I1 – I2 to give the approximate distance over which the boundary has shifted, Δx, between I_{base} and I_{reg}. This can be extended to three dimensions to approximate the volume loss, Δv.

8.2.2.5.Assessment and optimisation of the BSI

Optimisation of the BSI was performed on eight controls (average age 46 +/- 6 years) and eight genetically confirmed familial AD patients (average age 46 +/- 4 years) matched for age, sex and scan interval. These were a subset of the main group and their results are displayed separately in Figure 8.4, Figure 8.5 and Figure 8.6. The BSIs of this set of 32 hippocampi were calculated using ninety separate combinations of I_w and I_c. Agreement for each combination was assessed between the BSI and the difference between baseline and registered-repeat manual segmentations. A window width of 0.3 and centre of 0.65 (0.5 – 0.8) were chosen based on these analyses and these parameters were applied to all subsequent hippocampal BSI (HBSI) calculations. This window is higher and narrower than that usually used for whole brain BSI measures (a width and centre of 0.5 and 0.5 (0.25 – 0.75) (Freeborough and Fox, 1997)).

8.2.2.6.Assessment of location of BSI change measurement

In order to assess where the BSI measured change in the hippocampus a subgroup of baseline hippocampi were edited according to the following rules (see Figure 8.3). At a grey matter / CSF matter border, a line of voxels was removed and the regions were saved (GM-CSF negative regions). Loading the original hippocampus again, at the grey matter / white matter border, a line of grey matter voxels was removed and the edited region saved (GM-WM negative regions). Again the original hippocampus was loaded and the region was edited such that a line of CSF voxels was added to the hippocampus (GM-CSF positive regions); in a similar way, at the grey matter / white matter border a line of white matter voxels was added (GM-WM positive regions), and the regions were saved. The hippocampal BSI was then performed on these regions.

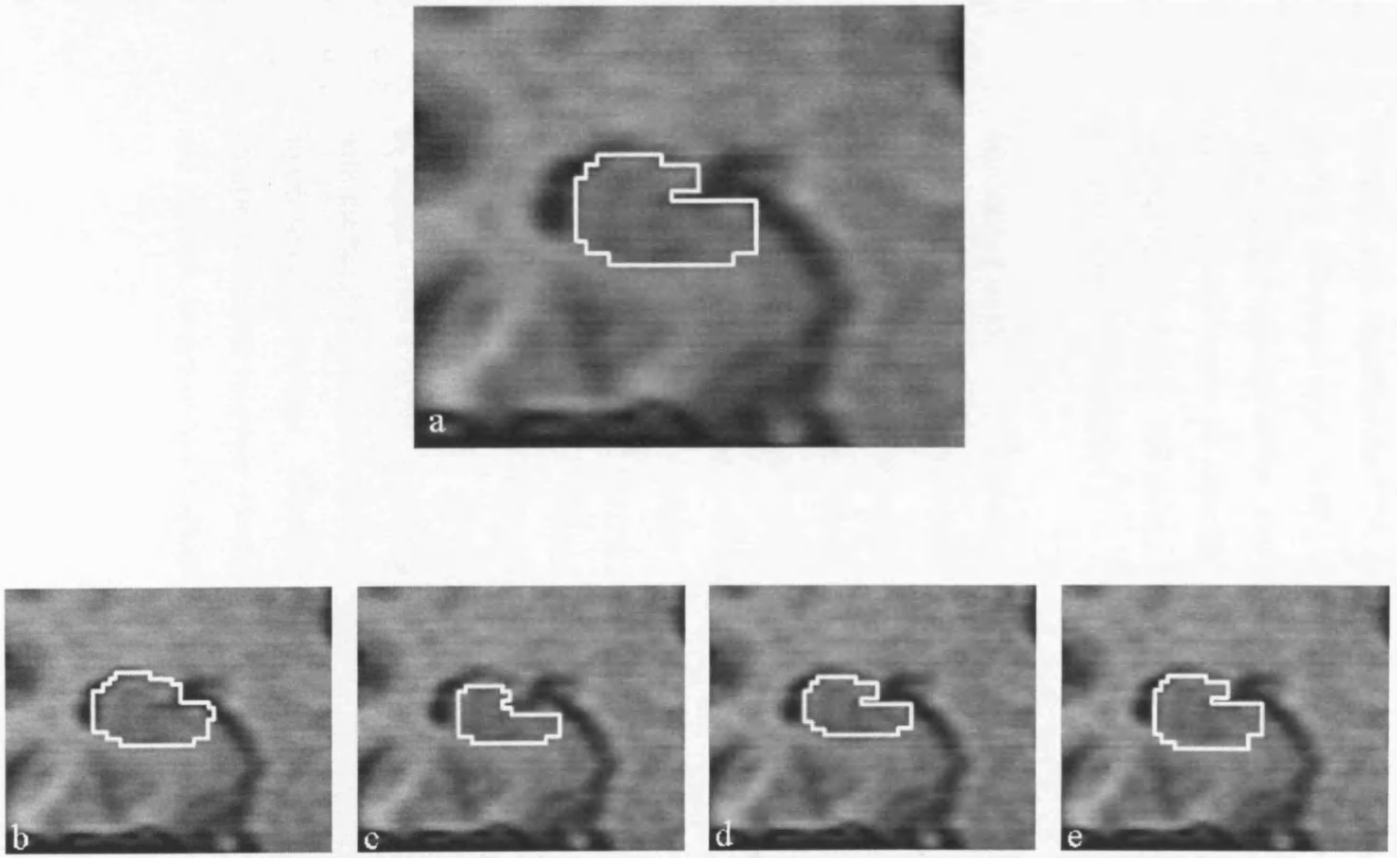


Figure 8.3 Edited hippocampal regions for experiment to assess location of border where BSI measures change in the hippocampus. a) is a standard hippocampus which is edited to generate: b) GM-CSF positive, where the hippocampal border was extended by one voxel at the grey matter- CSF border, c) GM-CSF negative, where a layer of voxels one voxel thick were removed at the grey matter – CSF border d) GM-WM negative, where the hippocampal border was extended by one voxel at the grey matter- white matter border, and e) GM-WM positive hippocampi, where a layer of voxels one voxel thick were removed at the grey matter – white matter border.

8.2.2.7.Statistical Analysis

Data were analysed using STATA version 8 (see Chapter 3, page 96). All hippocampal volumes were corrected for TIV, which was derived according to a previously described protocol (Whitwell *et al.*, 2001). The logs of the mean baseline and registered-repeat hippocampal volumes of the control group (n=36) were regressed against the logs of the mean TIV measures to establish the slope of the relationship between TIV and hippocampal volume. The resulting coefficient (α) was used to correct the TIV volumes as described in Equation 4.1.

Annualised rates of atrophy were calculated on the log scale assuming a constant proportionate rate of hippocampal loss. The relationship between TIV and rates of atrophy was investigated using logistic regression and as no evidence of a relationship was found, TIV was not included as a covariate in atrophy rate calculation. All hippocampal atrophy rates (BSI and manual) were calculated as left plus right hippocampus. Pitman's test was used to assess whether hippocampal BSI reduced the variation in hippocampal atrophy rates in AD and controls compared with the manual results. Specificities of the two techniques for a chosen sensitivity cut-off were calculated using a clinical diagnostic decision rule. Logistic regression was used to assess whether rates from hippocampal BSI were better than manually-derived rates in discriminating between the AD and control groups. Logistic regression was also used to assess whether atrophy rates derived from hippocampal BSI were in combination with the baseline volume better at discriminating between AD and controls compared to hippocampal BSI alone. Unpaired t tests (allowing for unequal variances in the two groups) were used to assess whether mean baseline hippocampal volume differed in AD and controls, both with and without TIV correction.

8.2.3. Results

Figure 8.4 shows the sum of left and right hippocampal volume at baseline with TIV correction. The means (+/- SD) of the AD and control groups were significantly different (AD 4568 mm3 (+/- 727) and controls 5405 mm3 (+/- 515) ($p< 0.001$)). However, there was a large degree of overlap between the AD and control groups. The mean difference (+/- standard error) between cases and controls with TIV correction was 838 mm3 (+/-149) and without TIV correction was 841 mm3 (+/-171). For a 90% sensitivity the specificity of baseline adjusted hippocampal volumes was 43%. Figure 8.5 shows the sum of left and right hippocampal volume at registered-repeat MRI (approximately 15 months later) with TIV correction. Again the means of the AD and control groups were significantly different (AD 4303 mm3 (+/-719) and controls 5350 mm3 (+/- 534)), but there was still a large degree of overlap. The mean difference (+/- standard error) between cases and controls was 1047 mm3 (+/-149) with TIV correction and 1057mm3 (+/-168) without correction. For a 90% sensitivity the specificity of registered-repeat adjusted hippocampal volumes was 53%.

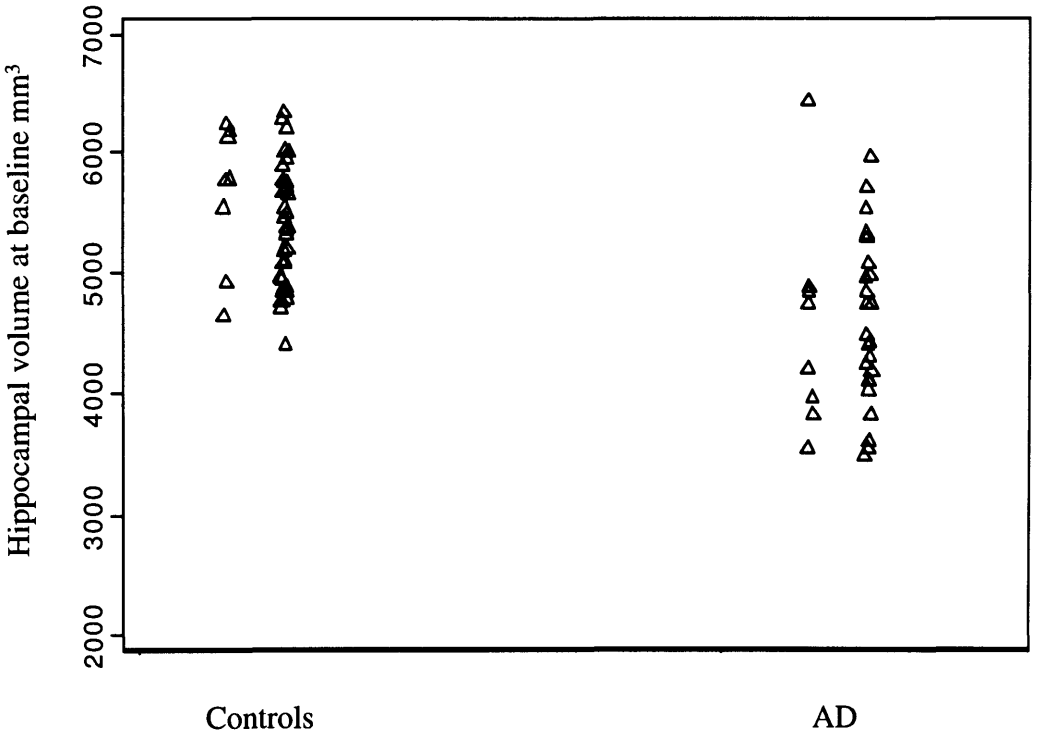


Figure 8.4 Manual measures of total (left plus right) hippocampal volume at baseline.
This figure shows the volumes of the manually outlined hippocampi at baseline with TIV correction. The optimisation set in both controls and patients are presented to the left alongside those from the whole group.

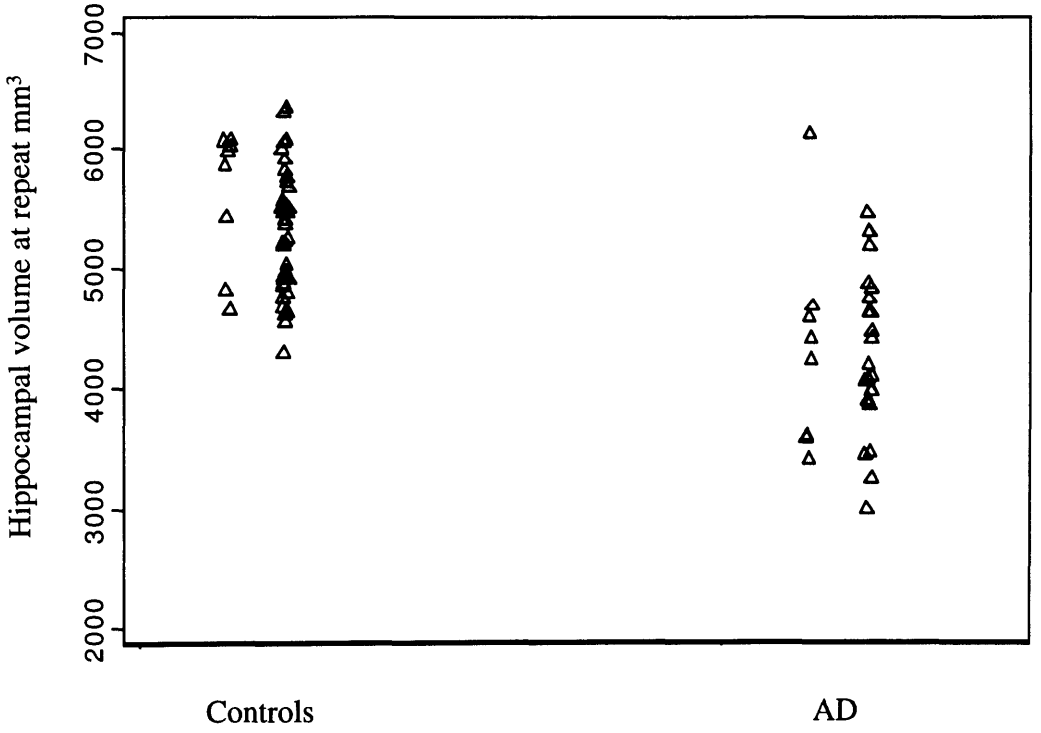


Figure 8.5 Manual measures of total (left plus right) hippocampal volume at repeat imaging.
This figure shows the volumes of the manually outlined hippocampi at repeat scanning with TIV correction. The optimisation set in both controls and patients are presented alongside those from the whole group.

.

Table 8.2 and Figure 8.6 show the rates of hippocampal atrophy in the total group of AD subjects compared with controls. In controls the variability in rates was significantly reduced using hippocampal BSI ($p < 0.001$). The mean rate was also significantly reduced ($p < 0.01$). In cases neither the mean rate, nor the variance differed significantly between the two methods of measurement.

Table 8.2 Mean (SD) rates of hippocampal atrophy, % year $^{-1}$ using both manual and hippocampal BSI techniques.

	Mean Rate (Manual)	**Mean Rate (BSI)**
Control	0.99 (1.88)	0.20 (1.06)
AD	5.77 (3.44)	5.86 (3.69)

This demonstrates rates of atrophy in both normal controls and AD subjects using both techniques. The optimisation set in both controls and patients are presented alongside those from the whole group.

Figure 8.7 shows the discriminatory ability of the two techniques. The sensitivity of the BSI-derived rates is greater than that of the manual rates at all specificity cut-offs largely because of the reduced HBSI rates in the control subjects. For a 90% sensitivity the BSI-derived rate had a specificity of 98% whereas the manually-derived rate had a specificity of 81%.

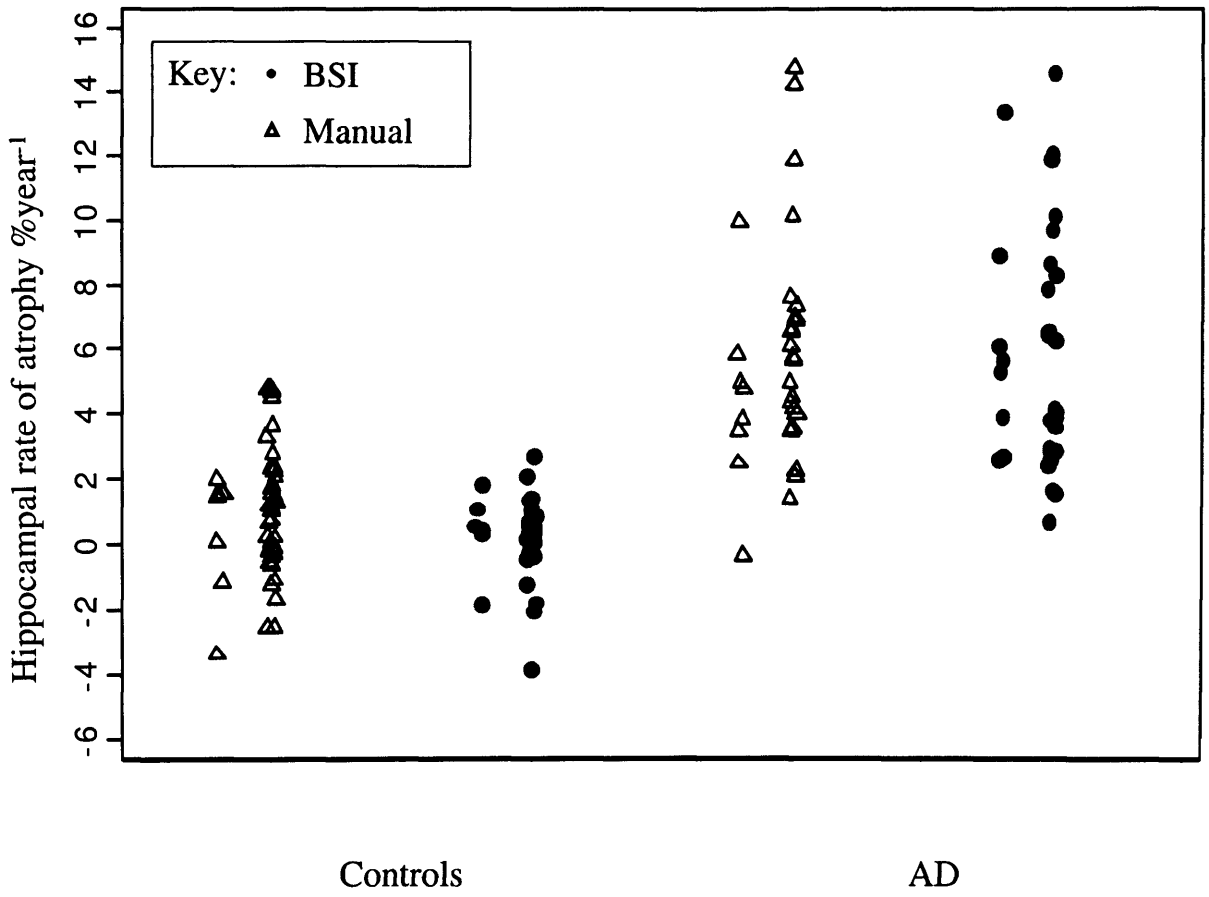


Figure 8.6 Manually-derived and BSI-derived rates of atrophy in control and AD subjects (% year $^{-1}$). Optimisation group is shown to the left alongside each group and method.

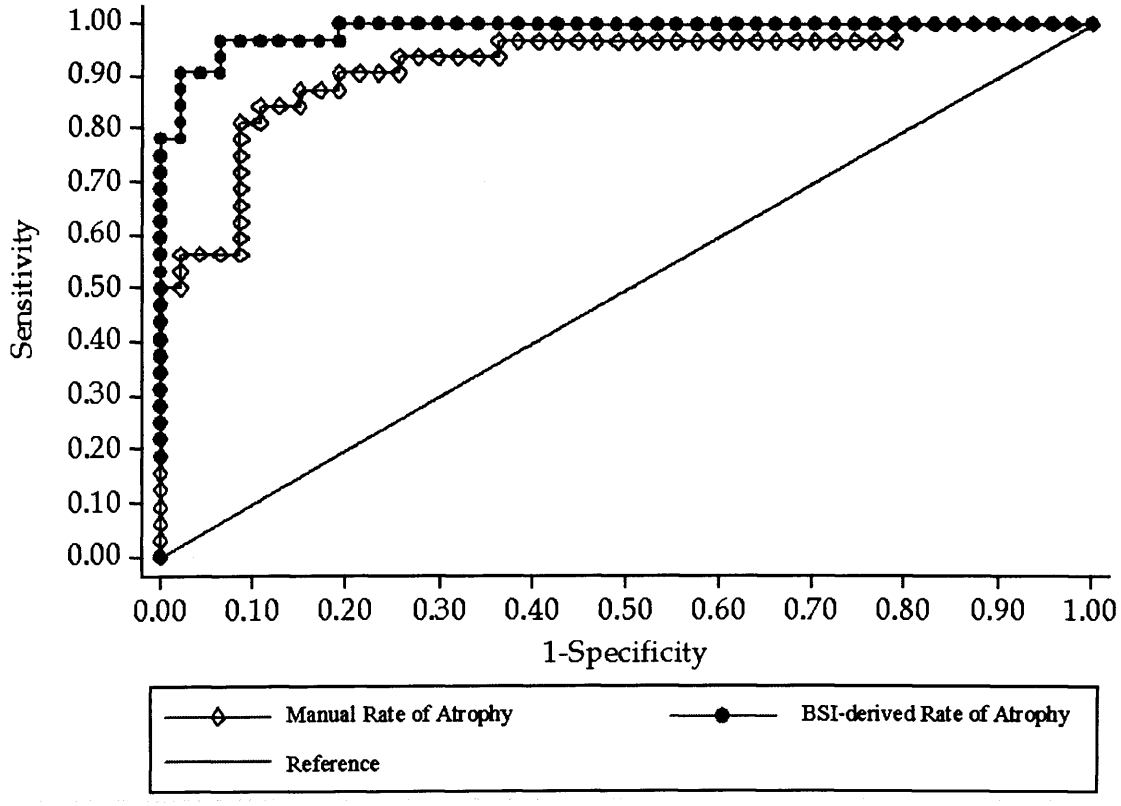


Figure 8.7 ROC graph of manually- and BSI-derived hippocampal atrophy rates.
This figure shows the sensitivity and specificity of both techniques in distinguishing between AD and control groups.

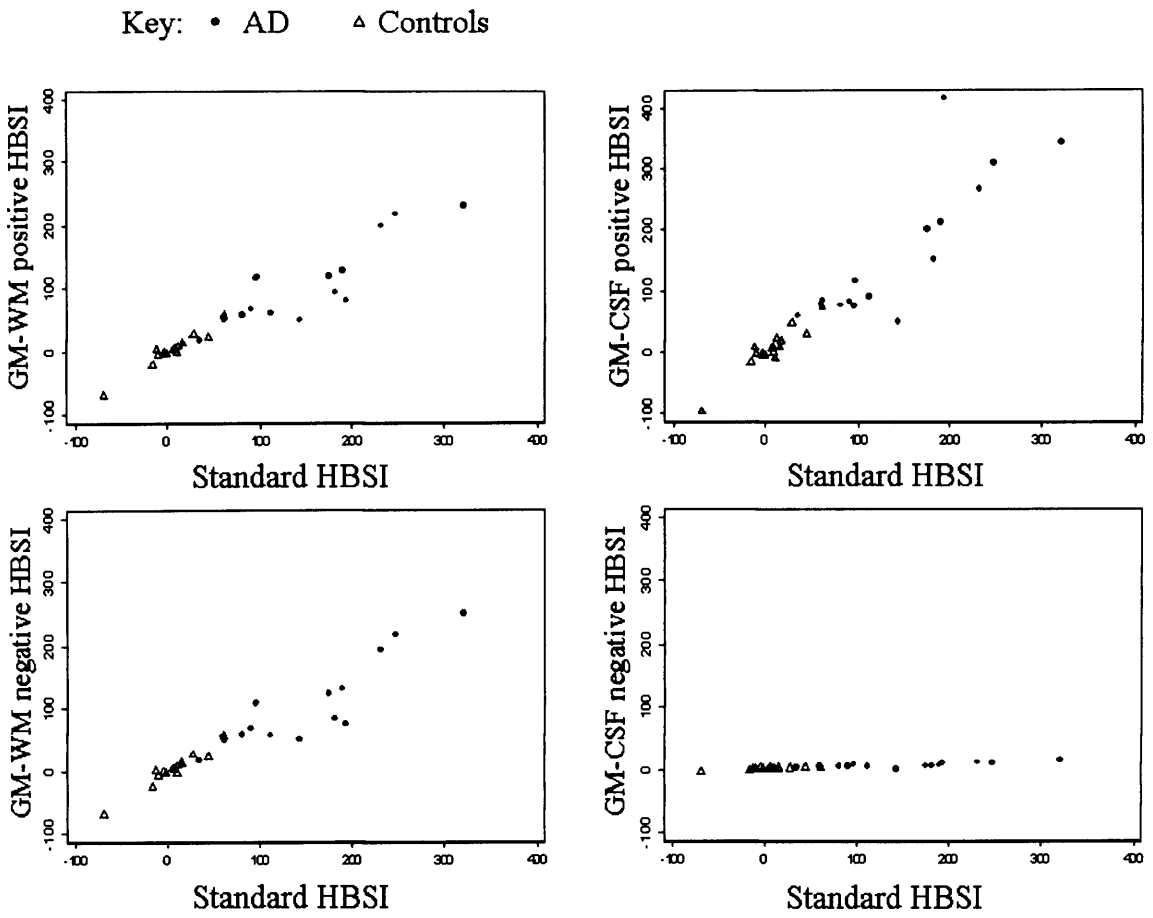


Figure 8.8 Scatterplots showing the associations of standard BSI measures against those from edited hippocampal regions. GM-WM positive HBSI (standard hippocampal region with addition of white matter voxels at white matter border). GM-CSF positive HBSI (standard hippocampal region with addition of CSF voxels at CSF border). GM-WM negative HBSI (standard hippocampal region with removal of grey matter voxels at white matter border). GM-CSF negative HBSI (standard hippocampal region with removal of grey matter voxels at CSF border).

From logistic regression models a 1% increase in BSI-derived rates of atrophy was associated with an 11 fold (CI 3, 36) increase in the odds of a diagnosis of AD. For manually-derived rates the equivalent odds ratio was 3 (CI 2, 4). Multiple logistic regression models showed that BSI-derived rates were significantly better than manually-derived rates in discriminating between cases and controls and provided no strong evidence that adding manual rates of atrophy to BSI-derived rates increased discriminatory power, for manually-derived rates ($p= 0.07$) and for BSI-derived rates ($p< 0.01$) in the multiple regression model. Furthermore, there was no strong evidence that adding adjusted baseline hippocampal volumes to the BSI-derived atrophy rates increased discriminatory power ($p= 0.09$).

Figure 8.8 a-d show the associations of hippocampal BSI calculated from the standard delineated regions (x axes on a-d) and from those from the edited standard regions (y axes on a-d). The edited HBSIs increased with standard HBSIs in an approximately linear fashion in all cases except that of the HBSI calculated using the CSF negative region (see Figure 8.8). This reflects the inability of the HBSI to calculate change in this case. This lack of ability to detect change was due to the fact that the area over which this atrophy was calculated did not include the grey matter - CSF border, and that change detected by the HBSI is calculated in this region.

8.2.4. Discussion and Conclusion

This study assessed the hippocampal BSI by comparison with the current "gold standard" of rates derived from manually outlined hippocampi on baseline and registered-repeat images. There was considerable overlap between control and AD groups in terms of hippocampal volumes (baseline and follow-up) even when these regions were corrected for differences in intra-cranial volumes. This concurs with previous studies that have shown considerable overlap between groups despite statistically significant group differences in hippocampal volumes (Convit *et al.*, 1993; Krasuski *et al.*, 1998; Mega *et al.*, 2002). In this study, adjusted baseline hippocampal volumes in the AD group were on average 20% smaller than those in the control group, but the diagnostic sensitivity and specificity of this measure was limited. A year later the follow up MRI showed the progression in hippocampal atrophy in the AD subjects had increased the difference between the groups but there was still considerable overlap between the AD and controls.

Rates of hippocampal atrophy had greater discriminatory power than either the baseline or follow-up hippocampal volumes. Mean rates of atrophy were six times higher in the AD group than in the control group (approximately 5-6% per year versus approximately 1% per year). The hippocampal BSI measure reduced overlap between the groups relative to the manual measurements by reducing both mean rate and variability in the control group. The HBSI is likely to underreport change as it primarily measures changes along the brain / CSF border. In the hippocampus, not all borders are brain / CSF. Large mis-registration or non-linear changes may also be a potential source of underreporting since the shifts in the CSF/hippocampal boundaries may not be fully captured. In addition, there is likely to be a bias in underreporting change between the control and AD groups. Relative to the AD group, the control group may have a lower proportion of their hippocampal border as CSF as less atrophy would have occurred before the first scan. Because of this, less change may be calculated along the hippocampal border. The control group is also likely to have much less intrinsic variability in hippocampal atrophy rates between subjects. Reducing measurement error will therefore make more of a difference to the variance in rates of atrophy in the control group than in the AD group where there is considerable between subject variability in disease progression.

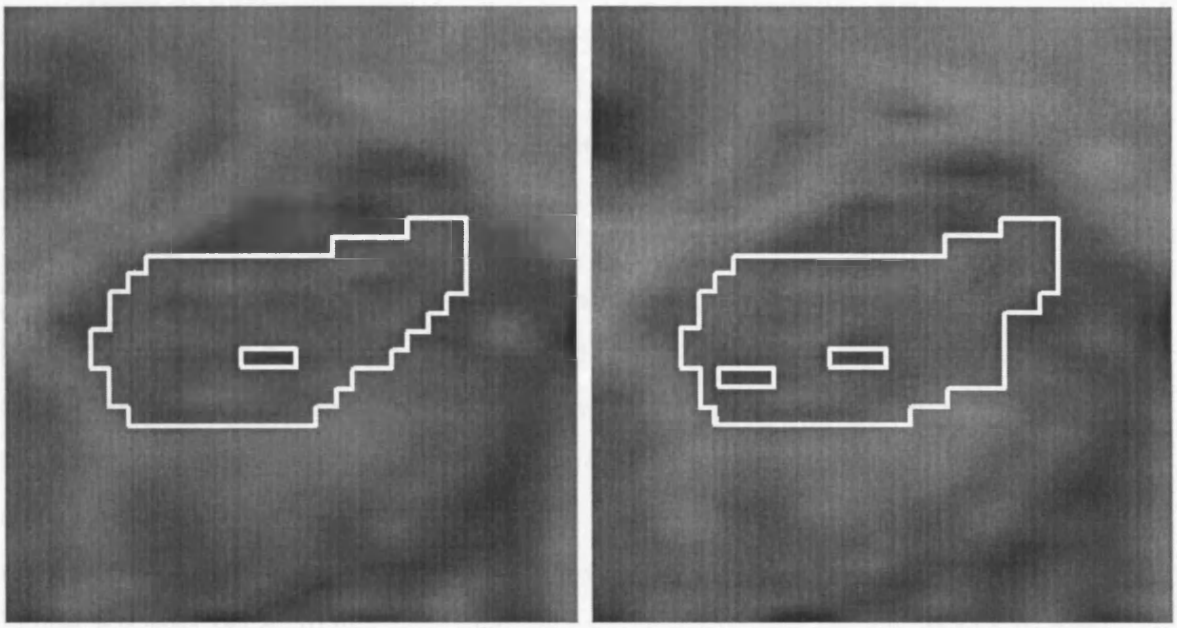
Rates of atrophy calculated by the BSI were 0.2% per year in the control group and 5.9% per year in the AD group. This was higher than quoted in the group of AD patients quoted in Chapter 5 (3.4% per year) (see page 143). This maybe due to the fact that the groups of patients were different, in this study the AD group was clinically-defined whereas those in Chapter 5 were post-mortem confirmed cases. In the PM-confirmed subjects, more than half of them were familial AD, whereas only just over one third of the subjects included in this study were familial AD. The differences seen may also be chance since the variance in the rates of atrophy measurements is high. However, the results from this study are consistent with the range of previously published rates of atrophy; a small study (n=4) showed early symptomatic familial AD rates were 4-8% per annum (Fox *et al.*, 1996b). Several studies by Jack and colleagues of older sporadic AD cases (mean ages 74-79 years in the different studies) have reported annualised rates of hippocampal atrophy of 3-4% with age-matched controls (mean ages 77-80 years) having rates of atrophy of 1.4-1.7% per year and similarly aged MCI subjects have intermediate rates at 1.8-3.7% per year (Jack *et al.*, 1998; Jack *et al.*, 2000; Jack *et al.*, 2004). Other groups have reported rates of atrophy in controls to be

0.8 % per year (mean ages 76 –77) and 5-6 % per year in AD subjects (mean ages 75-76) (Cardenas *et al.*, 2003; Du *et al.*, 2004). These differences may be due to differences in subjects (e.g. age or disease severity), in addition some differences may be ascribed to variation in segmentation protocols and differences in image quality and acquisition (Pruessner *et al.*, 2000).

The BSI is semi-automated and because it calculates change directly from differences between the MR images it is not so susceptible to errors of manual outlining. The BSI in the hippocampus calculates rates of atrophy by measuring the shift in the grey matter – CSF border and removal of hippocampal voxels at this border reduced the BSI measurement. Addition or subtraction of voxels along the grey matter – white matter border or addition of voxels along the grey matter – CSF border did not alter the calculation of the BSI greatly. The BSI has previously been shown to be a reliable and repeatable measure of whole brain atrophy in AD (Fox and Freeborough, 1997). The advantage of using the local hippocampal BSI is that it is possible to assess an area known to be susceptible to pathology at an early stage rather than the whole brain which includes large volumes that may not be affected by the disease at all or not until later in the course of the disease (Braak *et al.*, 1993; Braak and Braak, 1991; Scahill *et al.*, 2002). The BSI (or a variant thereof) has been applied to the general medial temporal lobe area in a normal-ageing cohort and was able to predict accurately 89% of those who would soon go on to decline cognitively (Rusinek *et al.*, 2003).

This study also demonstrates the existence of heterogeneity in rates of hippocampal change within the control and AD groups. This heterogeneity can only be partly explained by manual errors, misdiagnoses, or scan artefact; a significant component of the heterogeneity is due to true differences in progression of disease or natural ageing. Even with careful outlining manual errors can occur in a number of areas of the segmentation. Some error is inevitable in a small structure that varies in morphology between subjects and has a number of arbitrary cut offs and subjective judgements in outlining. Figure 8.9 shows two corresponding hippocampal "slices" manually delineated on baseline and registered-repeat scans of a control that yielded average results in both the BSI measure and manual measures. There are small differences in the manual outlining of arbitrary boundaries of the two hippocampi. This demonstrates that even with reference to another segmentation and registered image, small errors in delineation can occur. Other ways exist that are able to reduce the variance caused by

these errors in manual segmentation. These include repeating manual measures on the same images and calculating the average of the volumes for each hippocampus. Repeated measures should reduce errors however this can be very labour-intensive and in large studies or trials this may not be a practical solution.



BASELINE SCAN REPEAT SCAN

Figure 8.9 Manual segmentations of the hippocampus on baseline and registered-repeat scan.
These show two corresponding slices of hippocampi of a control subject at baseline and repeat imaging delineated using manual segmentation which gave average results using both manually-derived and BSI-derived techniques. This figure shows how small differences in segmentation (especially in the medial areas) can lead to noise in the manual measurements.

This study also established where the BSI was calculating change and confirmed that the BSI was measuring change at the CSF-GM border. This is somewhat unsurprising since it is known that the window width and centre (chosen by agreement with manual measures) were at levels which would predominantly measure this shift. It is this grey matter-CSF shift which is most obvious to the eye when flicking between baseline and repeat registered hippocampi. There is some evidence from other studies both volumetric (Csernansky *et al.*, 2000) and pathological studies (Von Gunten *et al.*, 2005), that it is the CA1 region (see Figure 8.10) which is most greatly affected by AD pathology. The CA1 region is a grey matter region located close to a CSF border in the lateral aspect of the hippocampus. As a result it may be the shift in this border that is predominantly measured by manual editing.

This study demonstrates how semi-automated measures of atrophy can potentially reduce within-subject manual errors, but Figure 8.6 shows that semi-automated techniques do not eradicate all errors in measurement. The negative values in Figure 3 represent potentially erroneous "hippocampal growth", and negative values were calculated using the HBSI measure. This may be due to errors in delineation, registration errors or quality differences between baseline and registered-repeat images. Figure 8.6 also demonstrates that semi-automated techniques do not totally remove overlap between the controls and AD. The remaining heterogeneity, present largely in the AD group, may be due to differing disease progression, differing involvement of the hippocampus, scan artefact, or between-subject delineation errors. In addition, only half of AD patients had the diagnosis of AD confirmed genetically or pathologically. In the other half of the AD group one cannot be sure that the underlying pathology is AD or purely AD (previous studies suggest diagnostic accuracy of around 90% (Galasko *et al.*, 1994)).

While the "gold standard" for measurement of atrophy of ROIs remains manual segmentation, various other methods have been proposed to assess atrophy rates either more easily and/or more precisely. Further research in these areas is needed. Non-linear registration may be used to propagate a baseline hippocampal region onto subsequent scans (Crum *et al.*, 2001). However, the clinical utility of this method, and susceptibility to scan artefact has not yet been assessed. Other techniques currently available require the manual selection of specific landmarks to guide non-linear registration of a template hippocampus onto the target hippocampus (Csernansky *et al.*,

1998; Haller *et al.*, 1996; Hsu *et al.*, 2002). Automation of segmentation is the natural goal for assessment of ROIs, however the manual intervention necessary for available "automated" techniques suggests a semi-automated approach is a reasonable step towards this goal.

In conclusion, rates of atrophy derived from serial MRI (two scans about one year apart) are superior to single measures of hippocampal volume in distinguishing between AD and controls. The BSI-derived rates of atrophy proved to be similar to manually-derived atrophy rates in terms of group discrimination. With only the baseline scans requiring hippocampal segmentation, application of the BSI method reduced both operator time and error, which is of particular importance when the hippocampus is to be used as an outcome measure in therapeutic trials.

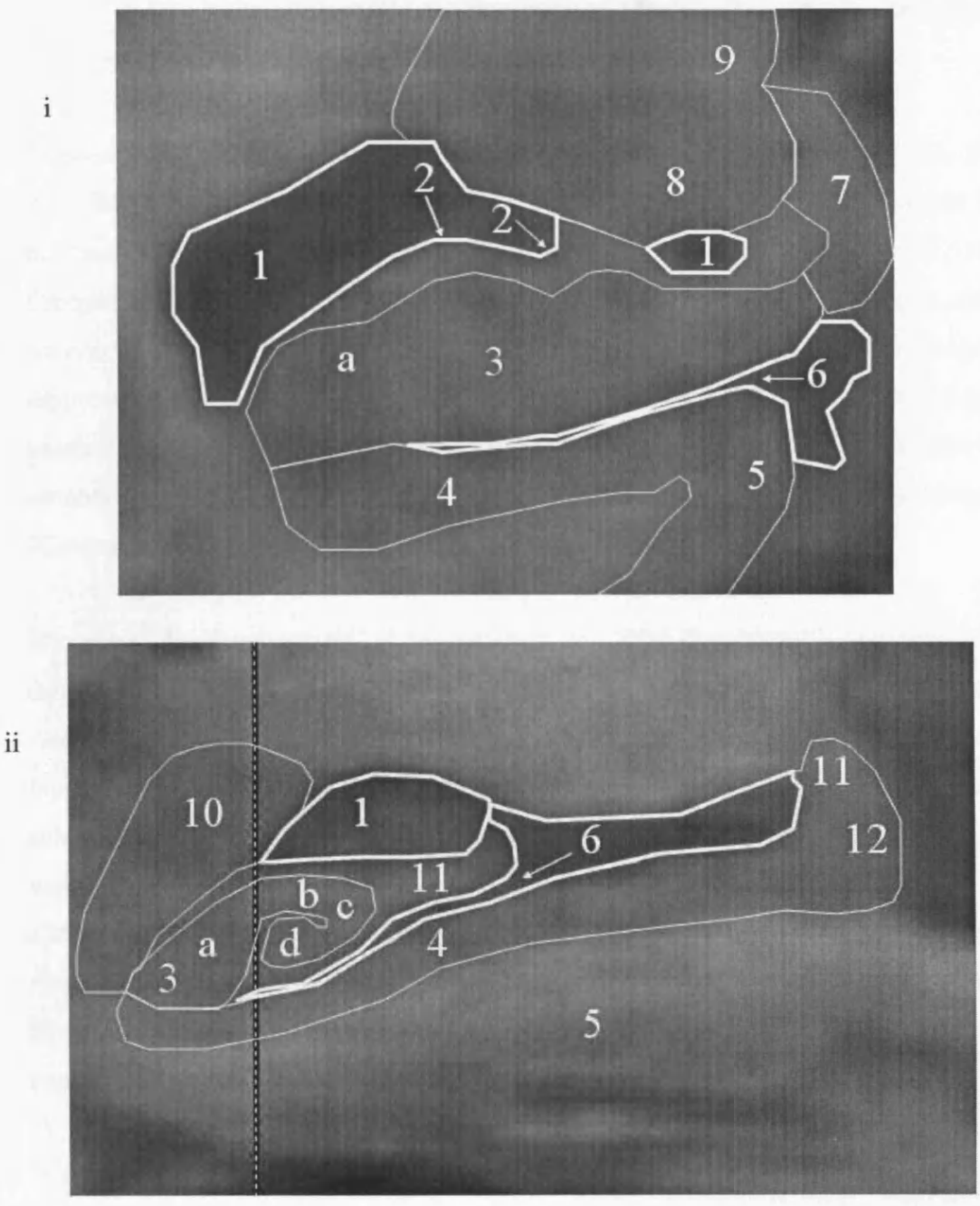


Figure 8.10 Schematic diagram overlaid on MRI image describing the anatomy of the hippocampus.
i) coronal view and ii) sagittal view of the hippocampus. Dashed line on ii) represents position of i) in the sagittal view. Key: (1) temporal horn of lateral ventricle;(2) internal digitations; (3) cornu ammonis(a CA1, b CA2, c CA3 and d CA4); (4) subiculum;(5) parahippocampal gyrus; (6) uncal sulcus; (7) subiculum in uncinate gyrus; (8) accessory basal nucleus of amygdala; (9) cortical nucleus of amygdala.

8.3. Automated measurement of hippocampal atrophy using fluid-registered serial MRI in AD and controls

8.3.1. Introduction

Various attempts at automated hippocampal segmentation have been made (see Table 2.1). Most require user intervention of some sort, and many of these techniques have not been tested in a clinical setting (Ashton *et al.*, 1997; Csernansky *et al.*, 1998; Csernansky *et al.*, 2000; Ghanei *et al.*, 1998; Gosche *et al.*, 2001; Haller *et al.*, 1996; Shen *et al.*, 2002). These methods have been developed mostly for the segmentation of hippocampal volumes at one time-point for volumetric group analyses. Only a small number of papers has applied an automated technique, which requires definition of a number of landmarks, to describe longitudinal volumetric changes in the hippocampus (Cardenas *et al.*, 2003; Du *et al.*, 2004; Wang *et al.*, 2003).

The term "fluid registration" (Christensen *et al.*, 1996; Freeborough and Fox, 1998) describes non-linear warping techniques based on the physical model of a compressible viscous fluid. These techniques make it possible to define a region (such as the hippocampus) on a baseline scan, and for this region to be fluidly-propagated through to subsequent serial scans. This can produce results at the same resolution as manual volumetry (i.e. a voxel is either "in" or "out" of the region of interest (ROI)). In addition, a sub-voxel level measure of change can be calculated by integrating the resulting Jacobian field. Fluid propagation methodology has been applied to the hippocampus, and has been shown to provide a reproducible measure with self-consistency better than a human rater (Crum *et al.*, 2001).

In this section fluid-based techniques are assessed, comparing atrophy rates at the sub-voxel and voxel-level resolutions and compare these results with those calculated from standard manual segmentation.

8.3.2. Methods

8.3.2.1.Subjects

Fifty-five controls and 32 AD subjects had two serial volumetric MRI assessments. The AD group comprised 14 FAD and 18 SAD cases; clinical details of these subjects are shown in Table 8.3. All subjects underwent clinical assessment including the MMSE (see Appendix Two) (Folstein *et al.*, 1975), and all patients fulfilled standard

NINCDS/ADRDA criteria (see Appendix One) (McKhann *et al.*, 1984) for the diagnosis of probable AD.

8.3.2.2.MR Imaging

T1-weighted volumetric MR brain scans were performed on a 1.5 Tesla Signa unit (General Electric, Milwaukee). The majority of scans (21 AD, 44 controls) were imaged using the standard protocol (see Appendix 3). Later scans (11 AD, 11 controls) were taken using the MIRIAD protocol (see Appendix 3). All volumetric scans were acquired as 124 contiguous 1.5mm coronal slices, which were transferred to a Sun workstation (Sun Microsystems Inc., Mountain View, Calif., USA) for analysis.

8.3.2.3.Region Segmentation

Brain regions were segmented on scans using a semi-automated iterative procedure using the software package MIDAS (Freeborough *et al.*, 1997). Prior to hippocampal segmentation, each scan pair (baseline and repeat) was co-registered with nine dof (rigid plus scaling) using the two segmented brain regions to accurately align the brains. Manual segmentation of the hippocampi was performed by two investigators according to previously described methodology (Scahill *et al.*, 2003; Watson *et al.*, 1992). Segmentations were performed using "mirror-image" volumetry (see page 101) such that all hippocampi were segmented on the right side of the image. The intra-rater reliability co-efficient (see Equation 2.3) for segmentation of the hippocampus was 0.98. Raters were blind to all patient information in addition to structural laterality and imaging time-point.

Table 8.3 Subject Demographics.

	Controls	AD
Number	55	32
M/F	27 / 28	13 /19
Mean (SD) age in years	57.1 (15.5)	59.1 (11.4)
Mean (SD) baseline MMSE	29 (1)	19 (5)
Sporadic/Familial	NA	18 / 14
Mean (SD) scan interval in days	449 (275)	450 (305)

NA = Not applicable

8.3.2.4.Fluid registration

Following the nine dof brain-brain registration an additional local rigid, six dof registration step was performed to optimise the match between the baseline- and repeat-scan hippocampi. This registration used the baseline hippocampus as a mask to limit the area over which the cost function was calculated.

The registered-repeat image was then fluidly registered to the baseline image. Details of the hippocampal fluid registration and propagation have been described previously (Crum *et al.*, 2001). The fluid registration was performed over a sub-volume of the images: a cuboid enclosing the hippocampal region as defined manually on the baseline scan. The repeat sub-volume image was warped to the baseline sub-volume image. The fluid algorithm iteratively drove the deformation field to maximise the cost function, which was the cross-correlation, whilst enforcing the deformation field to satisfy the fluid model. This ensured that the deformation field was diffeomorphic (invertible, differentiable).

The fluid equations were solved to give the velocity field, which was converted to displacement by multiplication with a suitable time-step. The time-step was set at each iteration to limit the maximum displacement within the hippocampal region. During the registration, trilinear interpolation was used. Once the exit criteria were met, 50 additional iterations were performed using sinc interpolation which provided further convergence to the optimal solution, but which was more computationally expensive than trilinear interpolation.

8.3.2.5.Choice of exit criteria

Potential exit criteria were assessed by observing the behaviour of various outputs of the fluid at each iteration for 2000 iterations. The potential candidates were the body force (the derivative of the cost function), the similarity of the fluidly-registered image and the baseline image (calculated as cross-correlation) and the amount of Jacobian change. These three were calculated over the whole hippocampus which was masked using the baseline hippocampal region. Exit criteria were initially assessed in a small group (two AD and two control subjects).

8.3.2.6.Fluid Propagation

The deformation field obtained from the fluid registration was then inverted to give the equivalent deformation field needed to deform the baseline to the repeat scan. The baseline hippocampal region was transformed by subdividing voxels in the repeat region and calculating from the inverse deformation field where each sub-voxel centre had displaced from the baseline image. For each voxel, if more than half of the sub-voxel centres had come from within the hippocampal region, it was included in the propagated region, otherwise it was excluded. Applying intensity thresholding to this binary propagated region ensured the hippocampus was labelled correctly on the repeat scan, and was consistent with the manual segmentation protocol. All fluid registrations and fluid-propagated regions were visually assessed.

8.3.2.7.Jacobian quantification

A field of Jacobian values was calculated from the deformation field obtained from the fluid registration. This gave the amount of contraction or expansion at each voxel. The amount of change within the hippocampal region was calculated at a sub-voxel level. This was achieved by integrating the Jacobian values within the baseline hippocampal mask.

8.3.2.8.Statistical analysis

Data were analysed using STATA version 8 and SAS (see Chapter 3, page 96). Atrophy rates were analysed for the total hippocampal region (left and right together) on a logarithmic scale [(log (follow-up volume/baseline volume))/interval] in order that doublings and halvings in volume be treated as effects of equal magnitude. Mean atrophy rates were calculated by back transformation with SDs calculated from variance transformation formulae. Comparisons between AD and control subjects were performed using t-tests allowing for unequal variances between the subject groups. Comparisons between methods for means and variances in the log atrophy rates were performed using paired t-tests and Pitman's test respectively. Sample sizes for randomised controlled clinical trials required to achieve a particular statistical power to detect a particular proportional reduction in an outcome variable in AD subjects are proportional to the square of the coefficient of variation (CV) of the outcome variable in cases. The relative numbers of patients required for two different outcome variables is therefore given by the square of the ratio of the respective CVs. 95% bootstrap confidence intervals (bias corrected) for such relative sample sizes (using a logarithmic transform for reasons of symmetry on a relative scale) were calculated using 10000 replicates.

Logistic regression was used to assess whether rates from any of the fluid measures were better than manually-derived rates in discriminating between the AD and control groups. ROC curves were generated to demonstrate graphically subject group separation according to technique. Sensitivities for 90% specificities cut-offs were calculated for each method. Linear regression was used to assess whether different scanning protocols influenced atrophy rates (adjusting for age) within each disease group and using each method.

8.3.3. Results

The results for running potential exit criteria are shown below in Figure 8.11-Figure 8.13 None of the output variables were stable enough to model the behaviour and establish non-arbitrary exit criteria. As result the body force was chosen and arbitrary criteria were established. These were satisfied when the mean body force ratio within the hippocampal region fell below $5.5*10^{-7}$. This was used for all resultant hippocampal pairs.

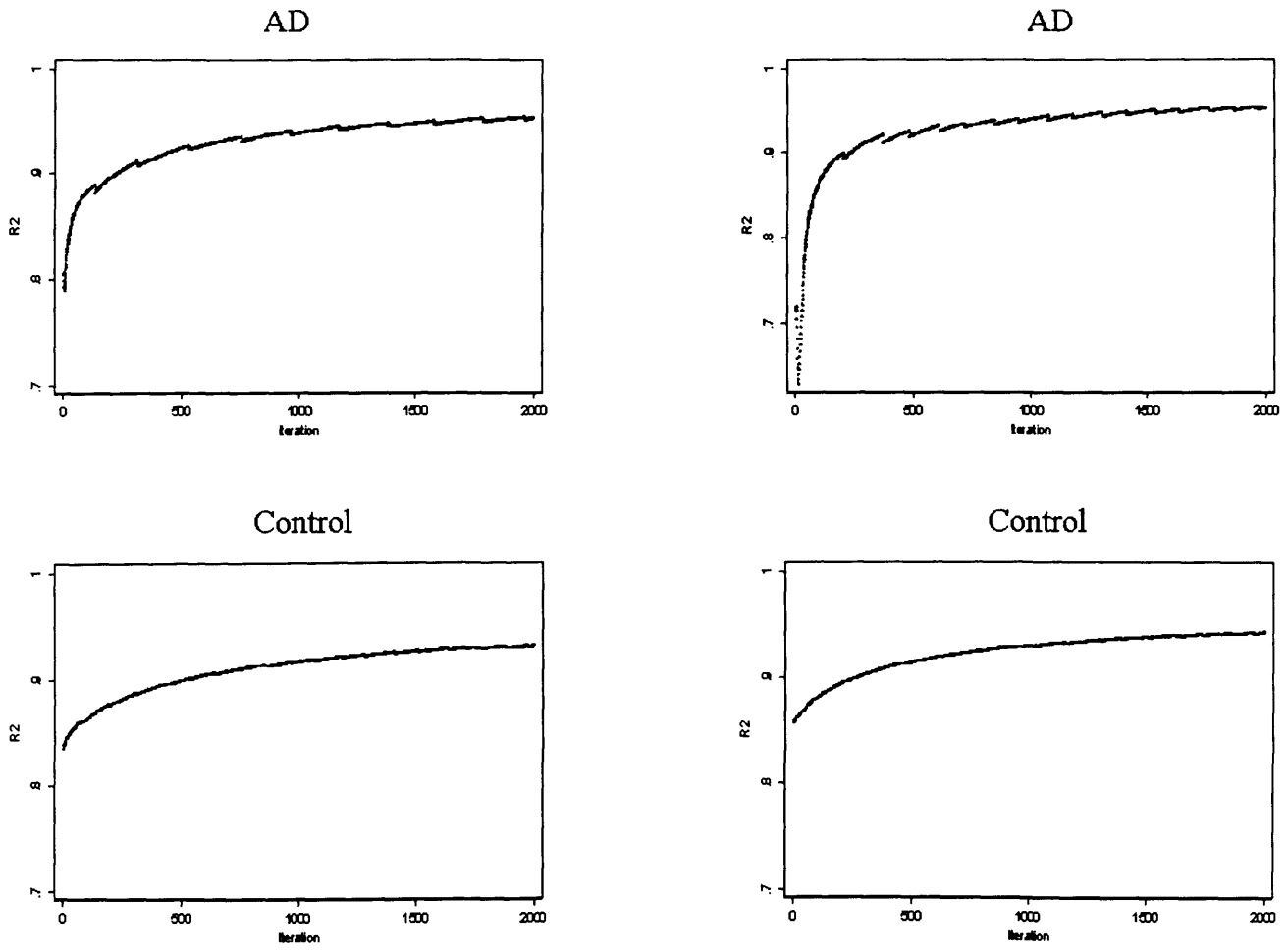


Figure 8.11 Potential exit criterion: R^2 of the baseline and fluidly generated image within the hippocampal region.

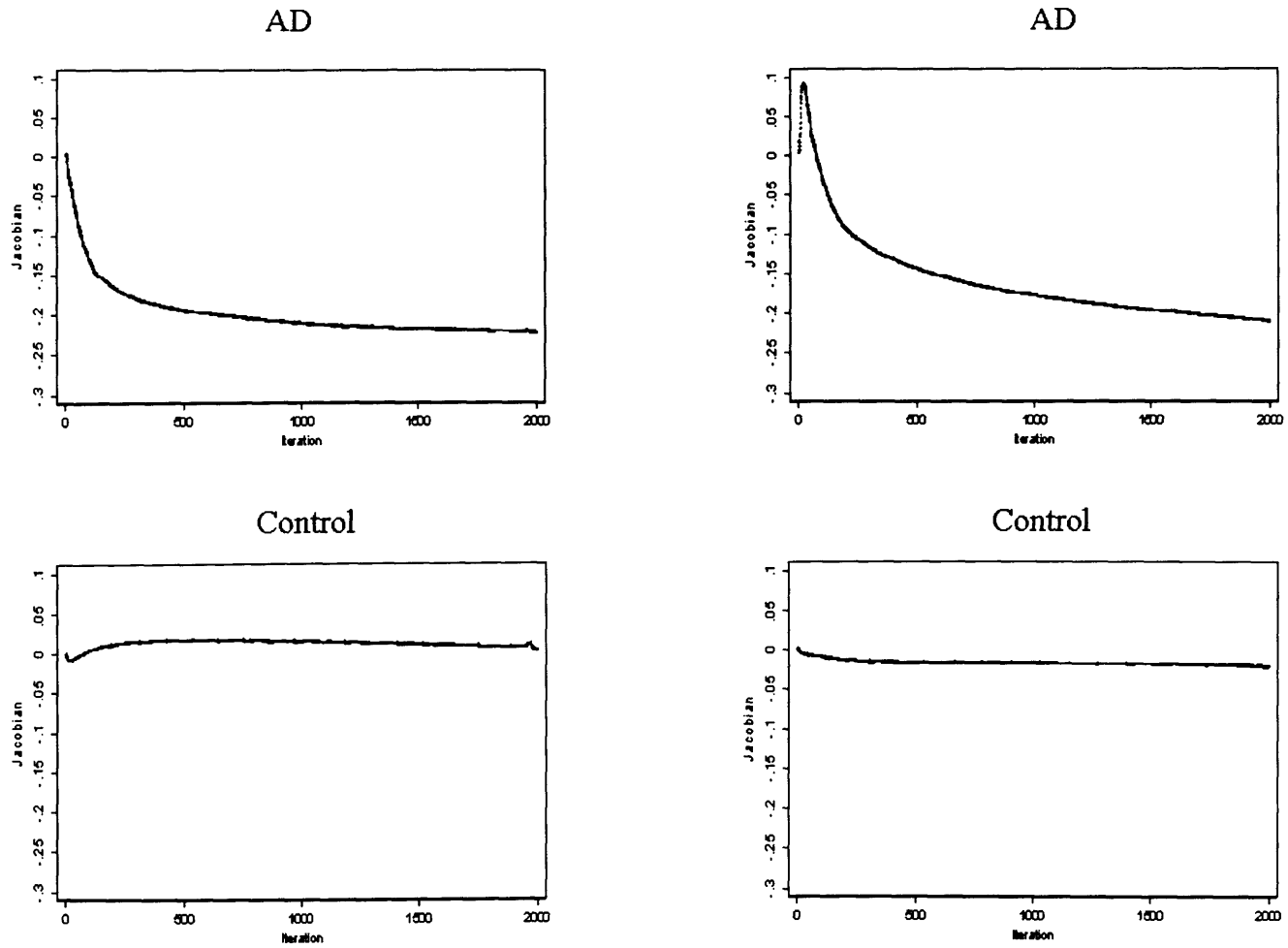


Figure 8.12 Potential exit criterion: Jacobian change between the baseline and repeat image within the hippocampal region.

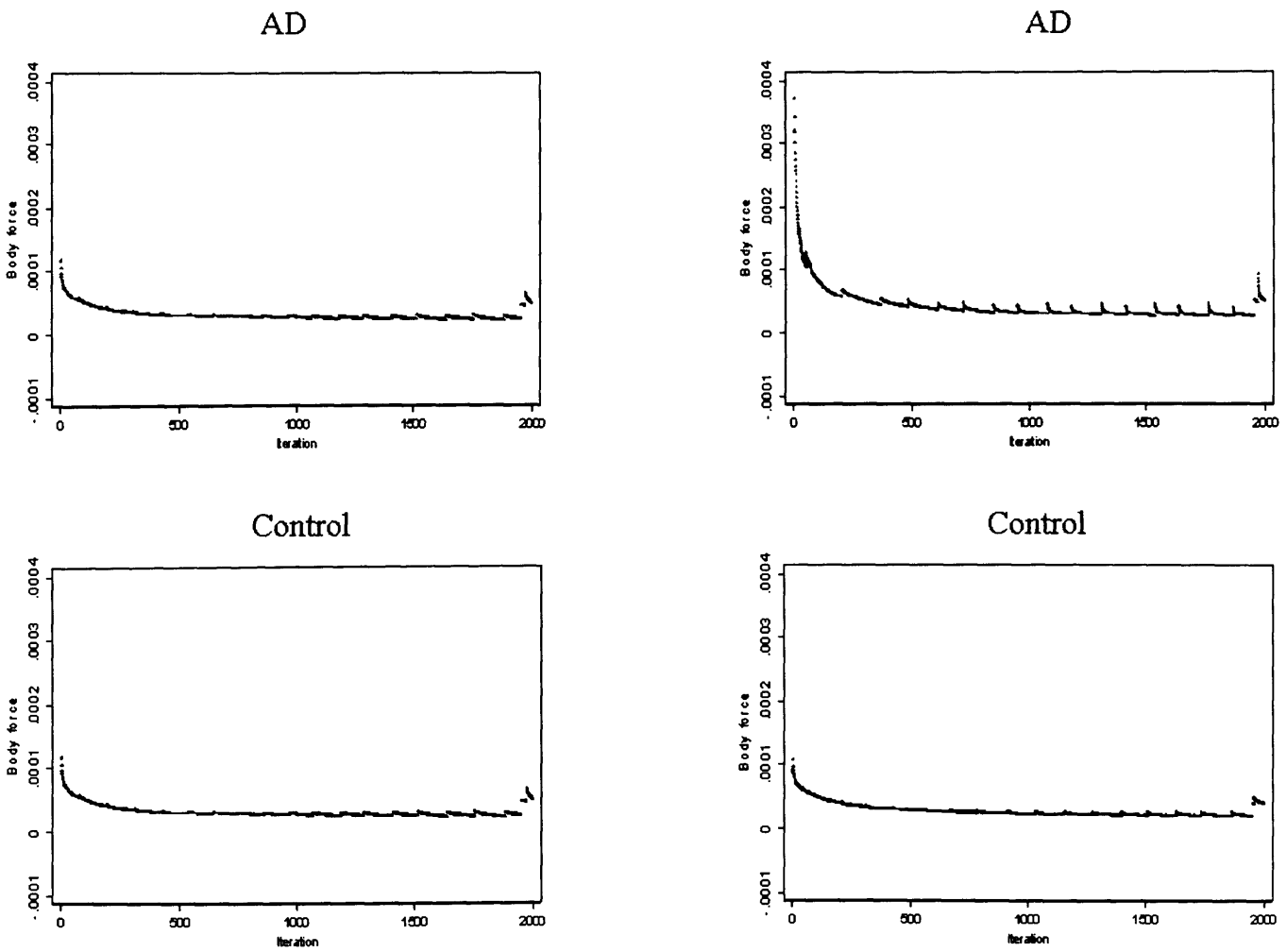


Figure 8.13 Potential exit criterion: Body force calculated within the hippocampal region.

Total hippocampal atrophy rates derived from manual, fluid-propagated regions and Jacobians are shown in Table 8.4 and Figure 8.14. Atrophy rates were significantly higher in AD patients compared with controls for all techniques (all $p<0.0001$). There was no evidence that different scanning protocols influenced atrophy rates within subject group using any method ($p>0.06$ for all comparisons). In AD patients there was no evidence that atrophy rates differed between the fluid propagation and manual rates in either means ($p=0.55$) or variances ($p=0.71$). However, there was evidence that rates using the Jacobian method were smaller in mean ($p=0.002$) and variance ($p=0.026$) compared with the manual rates.

In AD patients, the relative sample size for the fluid propagation method compared with manual atrophy rates was 0.83 but was not statistically significant (95% CI 0.48,1.42). The equivalent comparison of Jacobian rates with manual atrophy rates was also not statistically significant (relative sample size 1.14 (95% CIs 0.67,1.93)) reflecting the fact that both the mean and SDs were reduced using Jacobian rates.

In controls the mean atrophy rate from the fluid propagation was lower than that from the manual method but this difference was not statistically significant ($p=0.12$). There was strong evidence that control rates from the fluid propagation method were less variable ($p<0.0001$). As in AD patients, there was some evidence that Jacobian rates were smaller in mean ($p=0.014$) and less variable ($p<0.0001$) than the manual method.

Table 8.4 Mean atrophy rates (SDs) (%/year) for patients and controls derived from manual and fluid-based measures.

	Controls	**AD**
Manual volumetry	1.31 (2.00)	5.09 (3.59)
Fluid-propagated (with intensity thresholding)	0.89 (0.75)	5.34 (3.43)
Jacobian	0.56 (1.12)	3.55 (2.70)

All methods calculating atrophy rates were significant predictors of disease ($p<0.0001$). Figure 8.15 shows ROC curves for the total hippocampal rates using the three different techniques. For 91% specificity (50/55), the sensitivity of the manually-derived atrophy rate was 56% (18/32), and for fluid-propagated-derived rates this was 84% (27/32) and the Jacobian rates this was 72% (23/32). Multiple logistic regression analyses revealed that the combination of fluid-propagated atrophy rates and manual rates were better than manual alone ($p<0.001$) but not fluid-propagated alone ($p=0.679$), suggesting that the fluidly-propagated rate was a better discriminator.

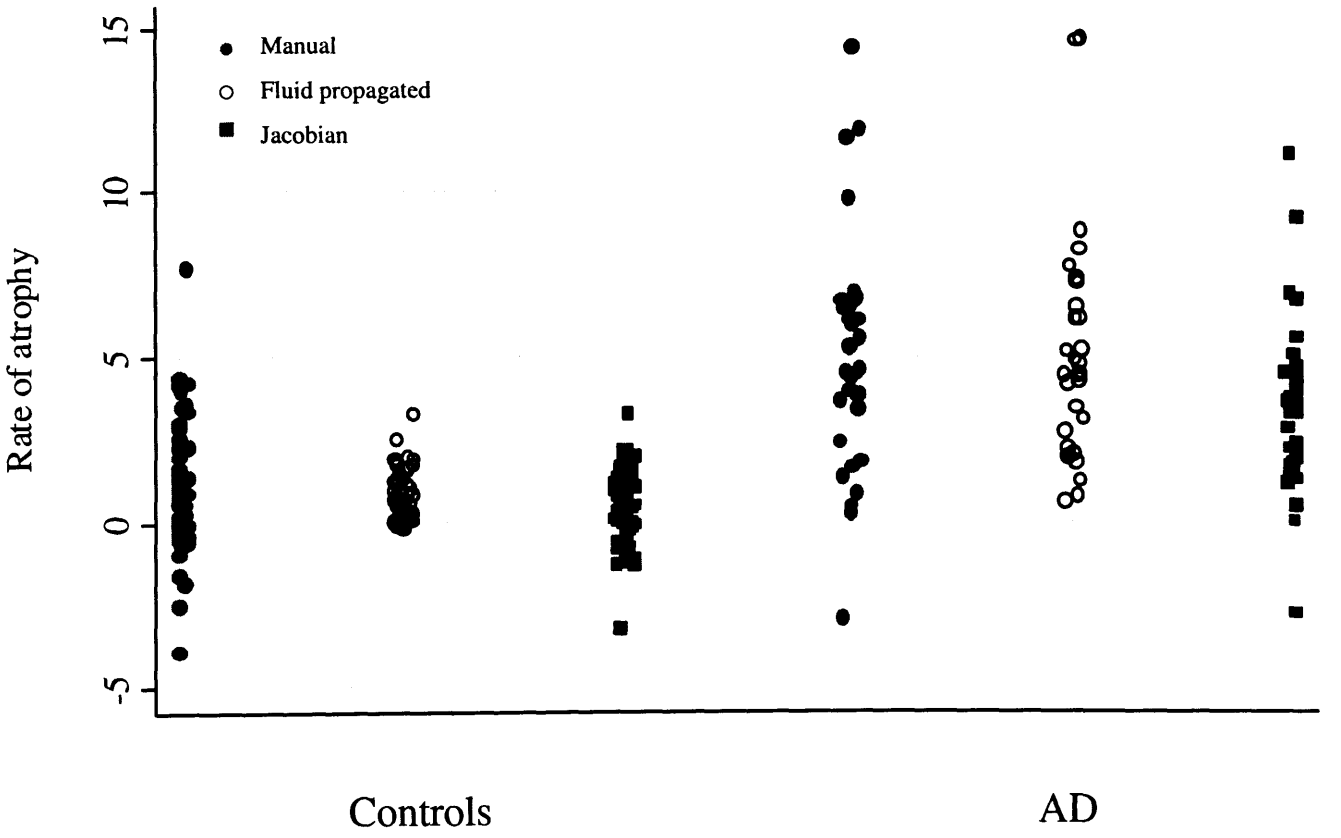


Figure 8.14 Rates of atrophy (left and right hippocampus, % per year) calculated from manual segmentations, and fluidly propagated regions with intensity thresholding (70% mean brain intensity) and Jacobian change.

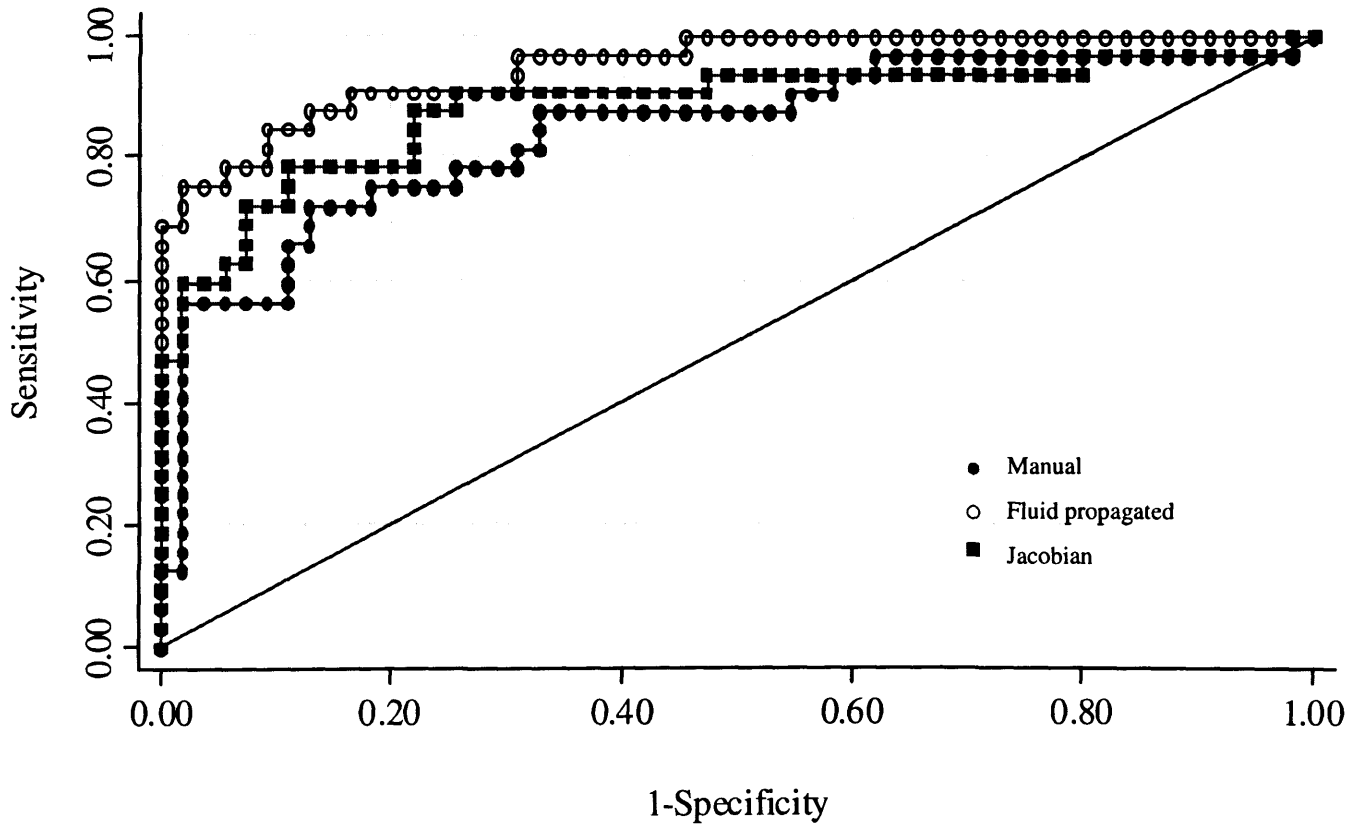


Figure 8.15 ROC curve showing group discrimination using manual, fluidly propagated and Jacobian rates of atrophy in the total (left plus right) hippocampus.

8.3.4. *Discussion*

In this chapter the ability of two measures of change calculated from hippocampal fluid registration to reproduce results similar to the manual measures and to discriminate between AD and controls was assessed. It was demonstrated that hippocampal rates derived from fluid propagation agree well with those derived from manual segmentation. The mean atrophy rates using the fluid propagation were slightly higher than manual atrophy rates in the AD subjects. This may be due to the fact that the intensity thresholding may also remove voxels from other areas of the hippocampus due to image artefact, real atrophy (i.e. the hippocampal fissure or uncal sulcus as it enlarges), or real biological change causing changes in signal intensity from the tissue which may not be volume loss (see Figure 8.16). One issue with such thresholding is that it is directionally biased as it will always remove voxels from the repeat region without including ones that may have been excluded from the first owing to artefact or noise. As a result, the atrophy rates calculated using such re-thresholding may be slight over-estimates in both subject groups. This study also showed that rates derived from fluid propagation were superior at separating AD subjects from controls than manually-derived rates. There was also a suggestion that fewer people would be needed within a trial compared with manual segmentation although this was not statistically significant and CIs around the relative sample size were wide.

The Jacobian measure of change is not subject to the arbitrary constraints of the voxel level measures (either intensity or voxel level binary thresholding). In theory, the Jacobian results should give the most accurate representation of change according to the fluid registration. Rates were found to be higher in AD than controls which was consistent with the voxel-level atrophy rates. The Jacobian measures were still significant predictors of subject group with manual measures in the same statistical model suggesting they each measure something different with respect to hippocampal change. However, Jacobian measures gave poorer separation between AD and controls than the fluidly-propagated atrophy rates. This may be due to the fact that real loss shown as small holes opening or appearing in the hippocampus will be missed by the Jacobian values as such holes will appear to have deformed from within the baseline hippocampus (see Figure 8.16). The Jacobian integration therefore will include the expansion of the hole itself, which will cancel out the compression of the surrounding area to leave a net change of zero. Small errors in baseline hippocampal delineation (for example including partial volume voxels, particularly in the uncal sulcus) may also

cause the loss calculated to be underestimated as these areas may expand between the baseline and repeat scans as the hippocampus decreases in size and this expansion may be included in the Jacobian value. Such delineation errors may lead to increases in the variance of Jacobian atrophy rates as well as the decrease of the means. The fluid-propagation circumvents this problem by including the intensity thresholding term.

This work builds upon a previous hippocampal fluid validation paper which described the technique for serial hippocampal region generation (Crum *et al.*, 2001). In that paper it was noted that where there was a great amount of change between scans of a pair, there was greatest mismatch between repeat manual and fluid-generated volumes. In this study the registration procedure was refined, including a further local rigid-body registration to optimise the matching of hippocampi prior to fluid registration and region propagation, which reduces the likelihood of mismatch influencing results. In addition exit criteria were applied to the fluid registration process so that each hippocampal pair finished when a similar level of matching had been achieved.

In this study, atrophy rates were significantly higher in patients with AD than controls, which was consistent with many previous studies (see Table 1.3). The atrophy rate calculated from Jacobian integration over the whole baseline hippocampal volume did have lower mean atrophy rates for both AD and control groups than is generally reported for manual studies for reasons described earlier (see Figure 8.16). The rates determined from the manual and fluidly-propagated methods were higher to those calculated in the post-mortem confirmed group described in Chapter 5 (see page 143), but Jacobian rates in this study were similar to the post-mortem rates. Discrepancies in mean rates calculated in different studies may be due to individual segmentation protocols employed by each research group, scan acquisitions, scan intervals or different ages and disease durations of cohorts being studied (Pruessner *et al.*, 2000).

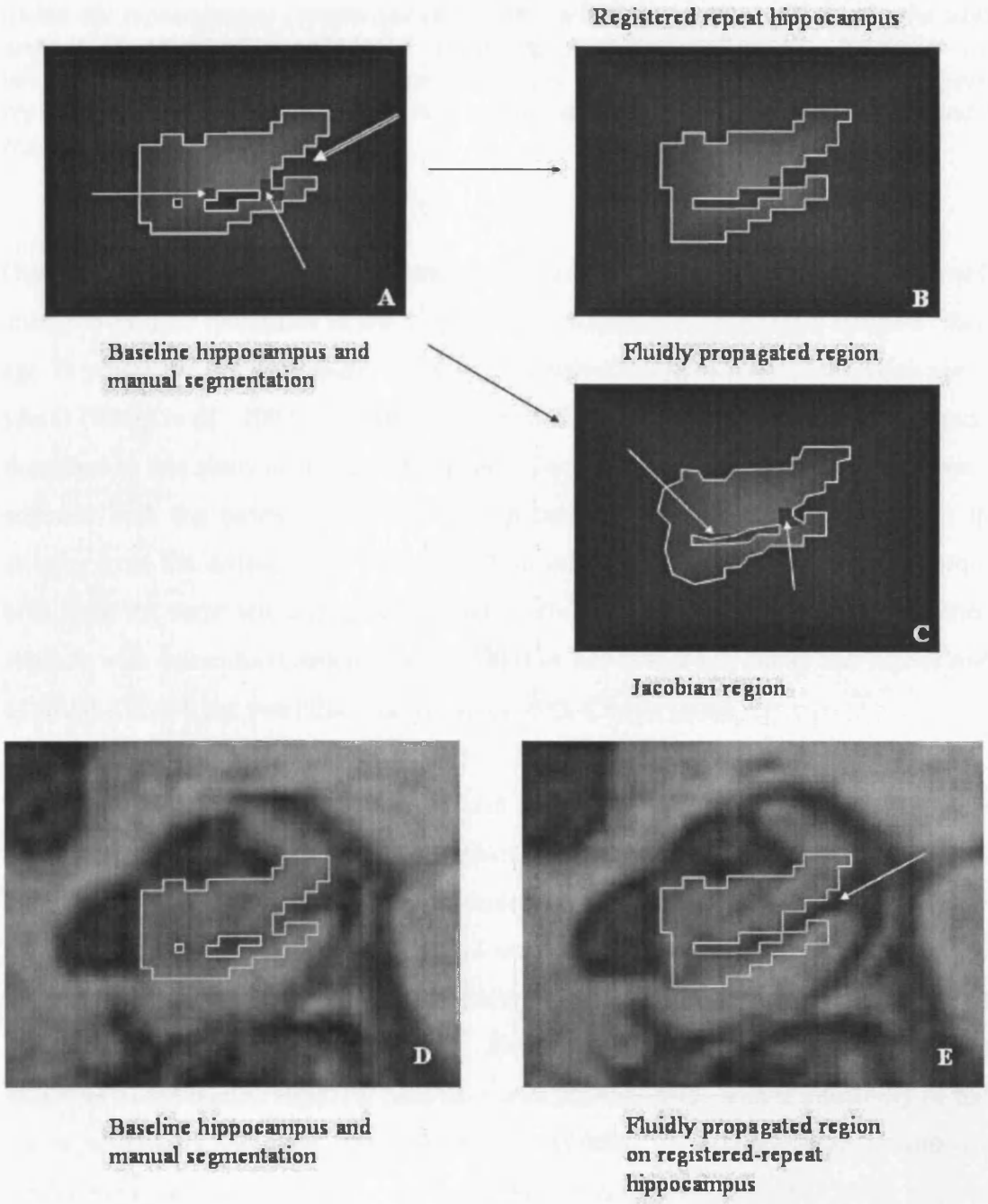


Figure 8.16 Schematic diagram demonstrating cause of potential differences in Jacobian and fluidly propagated rates.
Diagram showing coronal baseline (A) and registered-repeat (B) image of the right hippocampus of an AD patient demonstrating an increase in size of hippocampal fissure or uncal sulcus over the inter-scan interval (double arrow on figure A denotes fissure). Inter-scan interval is approximately one year. Figure A shows the manually-delineated baseline hippocampal ROI including some voxels which contain both hippocampal tissue and CSF (single line arrows); figure B shows the resulting fluidly-propagated region from the manually-delineated baseline. Figure C shows the region which may be labelled as repeat hippocampus according to the Jacobian method. Any partial volume uncal sulcus voxels included in the baseline segmentation (single line arrows on A) may be calculated as expansion by the Jacobian method of quantifying change and therefore these corresponding voxels on the registered repeat scan will be included

within the repeat region (single line arrow on C). This expansion will negate the same amount of contraction calculated within the hippocampus as the Jacobians are integrated over the baseline hippocampal region. The baseline and fluidly-propagated repeat hippocampal regions are shown on the actual registered MR images D and E respectively.

One study which utilised semi-automated non-linear registration measures to quantify change over time found loss in the hippocampus to be 2-2.5% in normal subjects (mean age 73 years) and 4-5% in mildly cognitively impaired / mild AD subjects (mean age 74 years) (Wang *et al.*, 2003). However, this technique differed from the fluid technique described in this study as it used a template hippocampus and non-linear registration to segment both the baseline and repeat hippocampi in all subjects and calculate the atrophy from the difference in volumes. Two other studies using a similar technique, both from the same research group, found similar results to Wang *et al.*, in that those subjects with dementia (Cardenas *et al.*, 2003) or AD (Du *et al.*, 2004) had higher rates of atrophy (5-6% per year) than control subjects (1-2% per year).

The fluid-based methods reported as part of this study gave good subject group separation. Many methods applied to structural imaging have been assessed for their ability to separate AD from controls. Within structural imaging, cross-sectional, semi-quantitative measurement of hippocampal atrophy has been shown to discriminate AD from controls with similar levels of sensitivity (81%) but slightly lower specificity (67%) than this study (Scheltens *et al.*, 1992). In addition, a combination of linear measures of the medial temporal lobe have also proved useful with a sensitivity of 85% for a specificity of 95% (Frisoni *et al.*, 1996b). Cross-sectional quantitative measurement of specific areas of the brain have also shown reasonable levels of group discrimination especially in the hippocampus and entorhinal cortex (Juottonen *et al.*, 1999; Killiany *et al.*, 2002; Pennanen *et al.*, 2004).

One strength of this study is the fact that there were relatively large groups of AD and control patients. Confirmation of the diagnosis of AD was available in 12 of the AD patients (two post-mortem and ten with proven pathogenic FAD mutations). A result of including FAD patients is that the mean age of the AD group was lower than that may be found in normal clinical practice. As atrophy of the brain (including hippocampus) increases with age (Scahill *et al.*, 2003), it is possible that the results reported may be

influenced by the relatively young age of the subjects and therefore this may need to be taken into consideration when relating results to an older cohort of subjects.

Manual segmentation, although the standard volumetric technique in current practice, is not a perfect measure of hippocampal volume. Arbitrary decisions are required to determine the boundaries of the hippocampus, and these can become particularly difficult to match on serial scans. Manual outlining may be open to bias, although an attempt was made to minimize this by presenting the scans in a random order with raters blinded to all scan and patient information. The exit criteria were based on the fluid achieving an arbitrary target (a body force ratio below a certain level). This convergence level and the intensity thresholding chosen in this study may need alteration when applying this methodology to separate cohorts with different scan protocols and qualities.

It can be concluded that fluid propagation has the potential to track volume changes in serial scans and is a valuable means of quantifying serial change in the hippocampus; with only the baseline hippocampus requiring segmentation, application of this technique reduces operator time. There may be benefits in using fluid propagation over manual segmentation of the hippocampus both for the purposes of diagnosis and monitoring disease progression. However, more work is required to establish whether this is the case in larger groups of subjects with differing clinical diagnoses, and more work is required technically to understand why the fluid does not seem to converge.

8.4. Chapter conclusions

Semi-automated measures of hippocampal atrophy are diagnostically useful and halve the amount of manual work performed by operators. In addition to this there is evidence that they can reduce the amount of operator error. As a result these measures may be useful in clinical trials and large studies where a large number of serial hippocampi may require segmentation or analysis of change within the region.

9. AUTOMATIC CALCULATION OF HIPPOCAMPAL ATROPHY RATES USING THE BOUNDARY SHIFT INTEGRAL

9.1. Chapter Introduction

Background

The development and validation of non- or minimally-invasive markers of disease is important in AD to aid diagnosis of the disease and to track disease progression. Early diagnosis is particularly important as it is those patients at the very earliest stages of the disease who may theoretically derive most benefit from disease-modifying therapies should they become available. For imaging markers to become useful diagnostically it is important that they become automated to reduce both the cost and labour. As described and discussed in Chapters 4 and 6 hippocampal volume has been shown with MRI to be reduced in AD subjects cross-sectionally although change in this volume obtained from serial scanning may have clinical advantages both for diagnosis and measurement of disease progression (see Chapters 5 and 8).

As described in Chapters 2,4,5, and 8, most volumetric studies of the hippocampus have used manual segmentation (Chan *et al.*, 2001b; Fox *et al.*, 1996b; Jack, Jr. *et al.*, 1992; Killiany *et al.*, 2002; Laakso *et al.*, 1995): this is currently considered to be the "gold standard" for measurement but is both time consuming and requires trained operators. Each hippocampus at each time-point takes approximately 45 minutes to delineate. In addition, manual outlining has a variable degree of operator error. As a result, efforts have been made to decrease the amount of operator interaction required for hippocampal segmentation. A number of semi-automated techniques have been developed (see Table 2.1). These may require manual outlining of the baseline hippocampus of a scan pair and measurement of shifts at the boundary of the hippocampus using the BSI (see Chapter 8, page 204) or fluid registration and integration of the voxel compression map to estimate atrophy in the hippocampal region (Crum *et al.*, 2001). Other methods require the definition of a number of landmarks on each scan to permit automated segmentation (Csernansky *et al.*, 2000; Duchesne *et al.*, 2002; Ghanei *et al.*, 1998; Gosche *et al.*, 2001; Haller *et al.*, 1997; Shen *et al.*, 2002; Wang *et al.*, 2003). One semi-automated technique has applied the BSI to measure longitudinal change in a region which approximates to the medial temporal lobe (Rusinek *et al.*, 2003). Further advances in automatically calculating hippocampal

volume or rates of atrophy may be of benefit both for diagnostic and clinical trial purposes.

Chapter Objective

The objective of the study described in this chapter was to develop an automated method of calculating hippocampal atrophy using inter and intra-subject registration and to compare this with both semi-automated hippocampal BSI methods and manual measurements.

9.2. Methods

9.2.1. Subjects

This study included 55 subjects: 36 clinically diagnosed AD patients and 19 age-matched healthy controls (one control and two AD subjects subsequently underwent autopsy with confirmation of diagnosis in all cases). Subjects were from the MIRIAD cohort (see section 3.3, page 92). All subjects underwent clinical assessment including the MMSE (Folstein *et al.*, 1975)(see Appendix Two), and all patients fulfilled standard NINCDS/ADRDA criteria (McKhann *et al.*, 1984) (see Appendix One) for the diagnosis of probable AD. Exclusion criteria were change in diagnosis at subsequent follow-up, or post-mortem confirmation of a differing underlying pathology. Subject demographics can be seen in Table 9.1.

9.2.2. Scanning protocol

T1-weighted volumetric MR brain scans were performed using the MIRIAD protocol (see Appendix Three).

Table 9.1 Subject Demographics.

	Controls	AD
Number	19	36
M/F	9/10	14/22
Mean (SD) age in years	68.7 (7.0)	69.6 (7.3)
Mean (SD) MMSE at baseline imaging	29.5 (0.7)	19.4 (4.1)
Mean (SD) scan interval in days	365 (5)	365 (18)

9.2.3. Manual Segmentations

All segmentations were performed using MIDAS software (Freeborough *et al.*, 1997). Brain regions on the baseline and repeat images were segmented using a previously described method (Freeborough *et al.*, 1997)(see Appendix Four), and all images were registered into MNI 305 atlas space (Mazziotta *et al.*, 1995) using six degrees of freedom (dof) to ensure all images were in a similar orientation. Following this the repeat images (second time-point) were registered to the baseline images using nine dof registration, to align the brains within a scan pair accurately. Hippocampi were delineated for each subject on the baseline and nine dof brain-brain registered-repeat images. The regions were delineated using "mirror-image" volumetry (see page 101). The baseline and registered repeat images were loaded simultaneously with the scans in a random chronological order. The hippocampus was always measured on the right-hand side of the presented image with the investigator blinded to the subject's name, diagnosis, chronological order, and left-right orientation of the scans. The protocol used is detailed in Appendix Four.

9.2.4. *Generation of standard automated single person template hippocampal BSI measures*

The standard automated hippocampal mask was generated using the following registration and quantification methodology (see Figure 9.1):

A single-subject template was chosen from the group of hippocampi (AD and controls). This 76 year old male control subject (MMSE 30/30) had average hippocampal volumes (manually-delineated) compared with the whole group.

1. Template to subject: Brain-Brain

An affine (12 dof) brain-brain registration of the single-subject template onto the baseline subject scan was performed (Woods *et al.*, 1998). The template hippocampal region was resliced using the obtained transformation parameters (all reslicing of regions of interest utilised trilinear interpolation with thresholding and interpolation of the images utilised Chirp Z (Rabiner *et al.*, 1969)).

2. Template to subject: Hippocampus-Hippocampus

Following this, an affine (12 dof) hippocampus-hippocampus registration of template scan to baseline subject scan was performed to align the hippocampi more precisely. The template hippocampal region was again resliced using the obtained local transformation parameters. This transformed template hippocampus was copied to the baseline subject image.

3. Subject follow-up to subject baseline: Brain-Brain

The repeat subject image used was that of the registered brain-brain (nine dof) to the baseline subject image (also used for the manual delineation). The repeat brain region and image were resliced using the obtained transformation parameters.

4. Subject follow-up to subject baseline: Hippocampus-Hippocampus

The repeat subject image was registered hippocampus-hippocampus (six dof) to the baseline subject brain using the transformed template hippocampus as the mask for the registration. The transformation parameters for this registration were applied to both the subject repeat brain image and region.

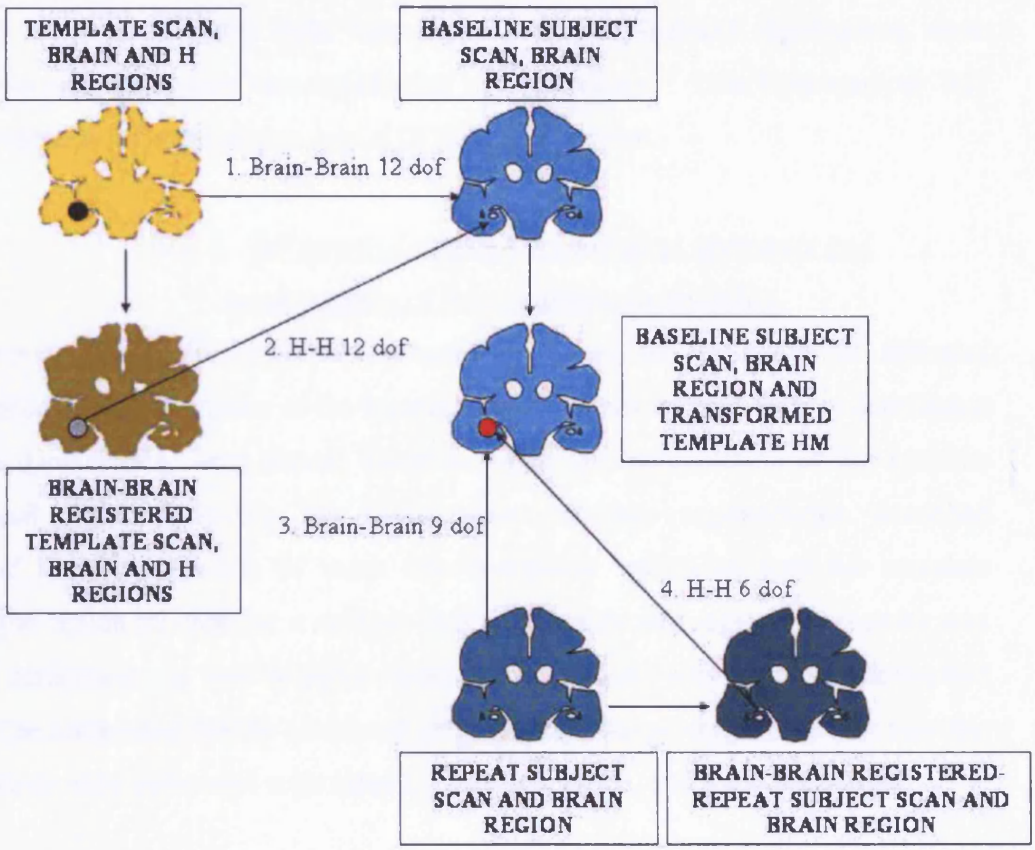


Figure 9.1 Registration methodology for generation of automated hippocampus masks. A template was chosen from the group of hippocampi (AD and controls). This individual had average hippocampal volumes (manually delineated) compared with the whole group. H = hippocampus, HM is the transformed template hippocampal mask.

The volume of hippocampal atrophy was then estimated by calculating the BSI (Freeborough and Fox, 1997) over volume *E* where:

$$E = HM \cap \left(d\left(b1 \cup b2\right) - e\left(b1 \cap b2\right)\right)$$

Equation 9.1

Where *HM* is transformed template hippocampal region, *b*1 is the baseline subject brain region, *b*2 is the transformed repeat subject brain region, *d* is a dilation by one voxel and *e* is the erosion by one voxel. *E* was dilated by two voxels to incorporate as many of the hippocampal boundaries as possible. This estimate is referred to as the "automated HBSI". The semi-automated HBSI was also calculated as previously described to allow comparison of measures (see Chapter 8, page 204). A window centre of 0.55 and a window width of 0.2 was used for the BSI in both the automated and

semi-automated measures (Freeborough and Fox, 1997). These procedures were completed for both left and right hippocampi. All intra-subject registrations were visually assessed to ensure mis-registration did not occur. This hippocampal BSI measure will be referred to as standard aHBSI from this point.

9.2.5. *Influence of simple morphological operators and incorporation of inter-individual variability*

In addition to the main experiment described above, the influence of different parameters on the segmentation of the baseline hippocampus and subsequent calculation of the standard aHBSI, were tested. These included the manipulation of the baseline hippocampal region following the single-person template segmentation described above, and the incorporation of some inter-individual variability into the template hippocampal region by creating a multiple subject template and regions. Accuracy was tested by comparing the new template-based segmentations with manually-delineated regions. The automated HBSIs calculated subsequent to the production of new baseline segmentations were compared with manual measures, HBSI, and standard aHBSI.

9.2.5.1.Assessment of different region manipulation schemes

Manipulations of single-person baseline regions, generated from the regions produced by trilinear interpolation, were achieved to establish whether the baseline segmentation accuracy could be improved by simple means. Baseline hippocampal regions created by the standard aHBSI single-person template were manipulated in a number of ways (see Figure 9.2) initially in a subset of the total cohort (three AD patients and three controls). The most accurate of these manipulations calculated using a similarity index (see Equation 2.2) was then performed in all subjects. HBSI was then performed using these manipulated regions instead of the single-person segmented template regions and the resulting rates of atrophy were converted to a rate of change to compare with the existing measures. From this point this will be termed manipulated aHBSI. There was one deviation from the standard aHBSI method which was that the hippocampal mask over which HBSI was calculated was only dilated by one voxel.

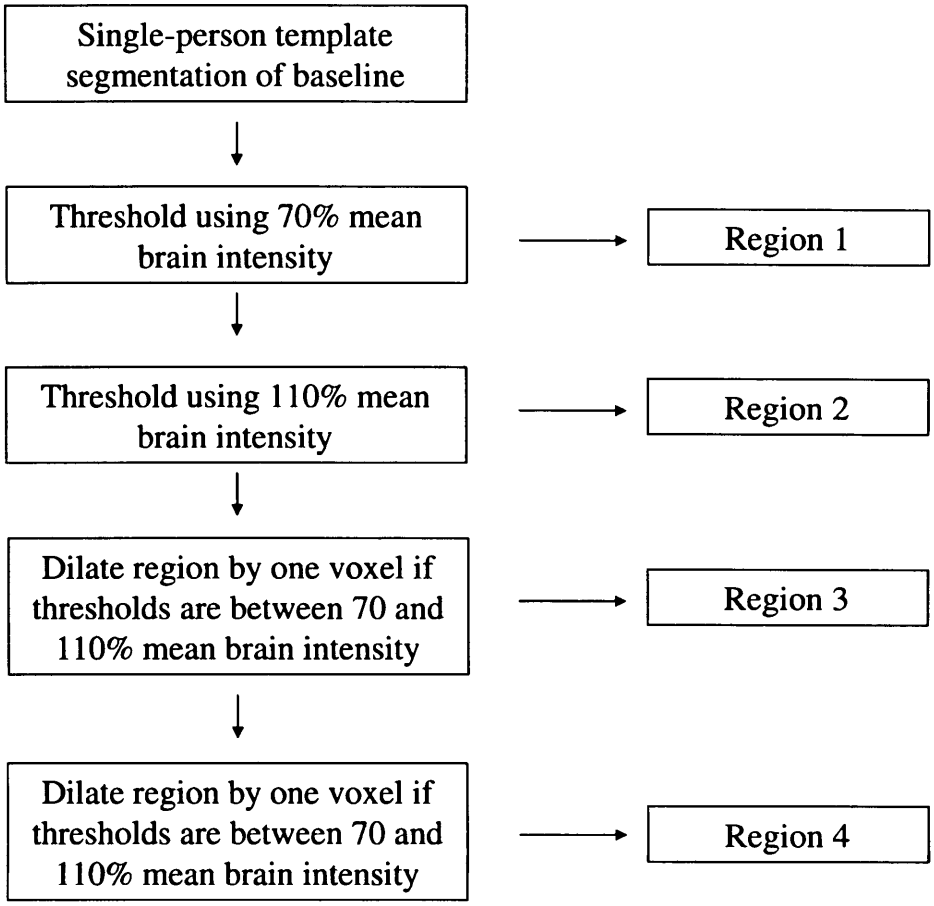


Figure 9.2 Manipulations of single-person template based segmentations. Regions were saved at each stage of this process allowing four new regions to be compared with the manually delineated baseline segmentations.

9.2.5.2.Assessment of average template hippocampal region

To assess whether the region in the template could be improved by incorporating inter-individual variability an average region was made. The average template region was produced by randomly assigning half of the AD and control subjects to a template generation group and half to a group on which the average template was to be assessed. This registration and template region production process can be seen in Figure 9.3. In brief, the scans from the template generation group were affine-registered brain to brain to the individual subject used as the single-person template previously (see page 246). The hippocampi were then rigidly-registered to this template and the binary regions were used to make a probability map where voxels had values between 0 and 1 (0 being no probability of containing hippocampus and 1 being all hippocampi occupy the voxel).

This map was then used to create regions (illustrated in Figure 9.3) according to this probability map: a region which encompassed all hippocampal regions (in other words, a region containing any hippocampal voxels in all scans) was generated and labelled "100% region", similarly regions were also created which encompassed 80% and 60% of all hippocampal voxels in all scans were created and labelled "80%" and "60% regions". These three new regions were used in generating new baseline hippocampal regions and subsequently the HBSI as described in Figure 9.1 in the remaining half of subjects. From this point this is named multi-subject aHBSI. Again, there was one deviation from the standard aHBSI method which was that the hippocampal mask over which HBSI was calculated was only dilated by one voxel.

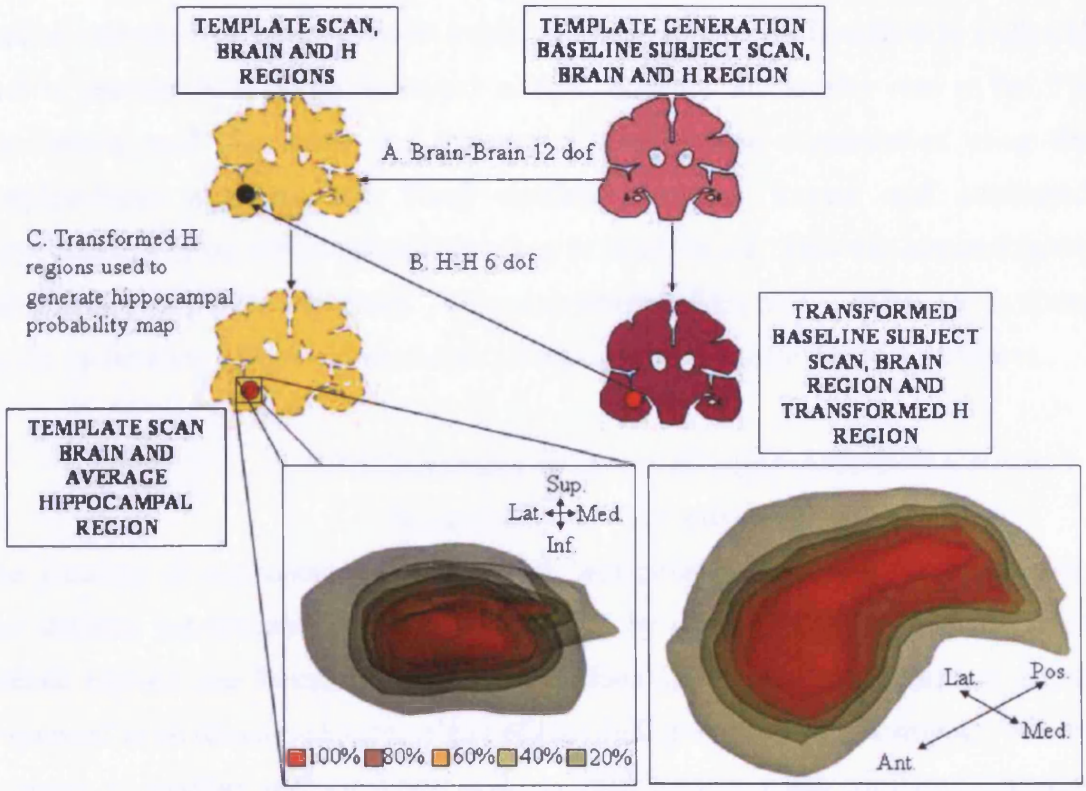


Figure 9.3 Schematic diagram showing production of multi-subject hippocampal template regions.
Half of the AD and control subjects (purple brains) were registered to the single-person template image used previously. This registration was performed A. brain to brain and B. hippocampus to hippocampus. Hippocampal regions were resliced at both stages and the transformed regions were summed producing a probability map which enabled the creation of probability-based regions (shown in the inset boxes). These are smoothed and rendered regions and percentages relate to the level at which voxels encompass the hippocampal regions in all subjects i.e. all hippocampi fall within the 100% region. These created regions were used in the generation of multi-subject aHBSIs by feeding the regions into the automated HBSI algorithm shown in Figure 9.1. Key: H = hippocampus, Sup. = superior, Inf. = inferior, Med. = medial, Lat. = lateral, Ant. = anterior, Pos. = posterior.

9.2.6. Statistical Analysis

9.2.6.1.Standard aHBSI single-person template analysis

HBSI and aHBSI were applied to each image pair and the resulting quantification was expressed as an annualised volume loss of total (left plus right) hippocampus. Equivalent measures for manual atrophy rates were calculated in the same way. Means of atrophy rates between AD patients and controls for each method were compared using t tests allowing for unequal variances. The means calculated by the different methods were compared using paired t tests. Differences in variances between the

methods were assessed using Pitman's tests. Each method's ability to separate AD subjects from controls was assessed using logistic regression. Power calculations were used to estimate how many subjects would be required using each method in a clinical trial to provide 90% power to detect a 20% reduction in atrophy rate at the 5% significance level. To assess the accuracy of the baseline segmentation using the template-based measures, the voxel similarity between manual and automated hippocampal regions was calculated described in Equation 2.2. This was assessed in the AD and control groups separately. To assess whether there was a difference in voxel similarity between AD and controls t tests were used allowing for unequal variances.

9.2.6.2.Analyzing influences of region manipulation schemes incorporation of multi-subject variability

The accuracy of segmentations generated by manipulating aspects of the registration and template segmentation process was assessed by comparing voxel similarity with manual regions (see Equation 2.2). The resulting HBSIs were calculated as above (expressed as an annualised volume loss of total (left plus right) hippocampus). Where accuracy of baseline segmentations were reasonable (>0.6 voxel similarity), relative annualised hippocampal rates of atrophy were also calculated. Paired t tests were performed separately in AD and control groups to establish whether the altered automated baseline segmentations were significantly more accurate than the standard automated segmentation. Manipulated aHBSIs generated from these altered regions were compared with standard aHBSI and semi-automated HBSI as well as manually-derived rates to assess whether any improvement of accuracy of segmentation was reflected in an improvement in association with the more validated measures. Similar analyses were performed with the multi-subject aHBSI measures.

9.3. Results

9.3.1. Standard single person template-based HBSI

Table 9.2 shows the mean (SD) annualised rates of total (left plus right) hippocampal atrophy for manual, HBSI and standard aHBSI, and the voxel similarity for the automated region. Rates were calculated on an absolute scale since the accuracy of the baseline regions generated from the template were relatively low. Figure 9.4 shows individual results using each method in control and AD groups. There was evidence of a difference in rates of atrophy between the AD and control groups ($p < 0.001$ for each method). There was no evidence of a difference in means between HBSI and manual

methods (p = 0.33) or standard aHBSI and manual methods (p = 0.75). SDs were comparable using HBSI, standard aHBSI and manual measures. Differences did not reach significance for any comparison (p > 0.2).

There was generally good individual agreement between HBSI and standard aHBSI methods, with the SD of the difference between the methods being 36.61 mm^3. Such agreement was not as good when comparing manual and HBSI, and manual and standard aHBSI with the SDs of the differences between methods being 82.77 and 91.28 mm^3 respectively.

Table 9.2 Annualised hippocampal atrophy rates (mm3 loss /year) (mean (SD)) using manual, HBSI, and standard aHBSI measures.

	Manual	**HBSI**	**aHBSI**	**Voxel Similarity of aHBSI region to manual region**
Level of manual outlining	Baseline and follow-up	Baseline only	Template only	
Controls	18.1 (53.5)	15.3 (50.2)	11.3 (50.4)	0.49 (0.05)
AD	174.6 (106.5)	159.4 (101.2)	172.1 (123.1)	0.43 (0.07)

Manual (outlining of both baseline and repeat scans), HBSI (outlining just the baseline scan), and standard aHBSI (fully automated template-based BSI) measures.

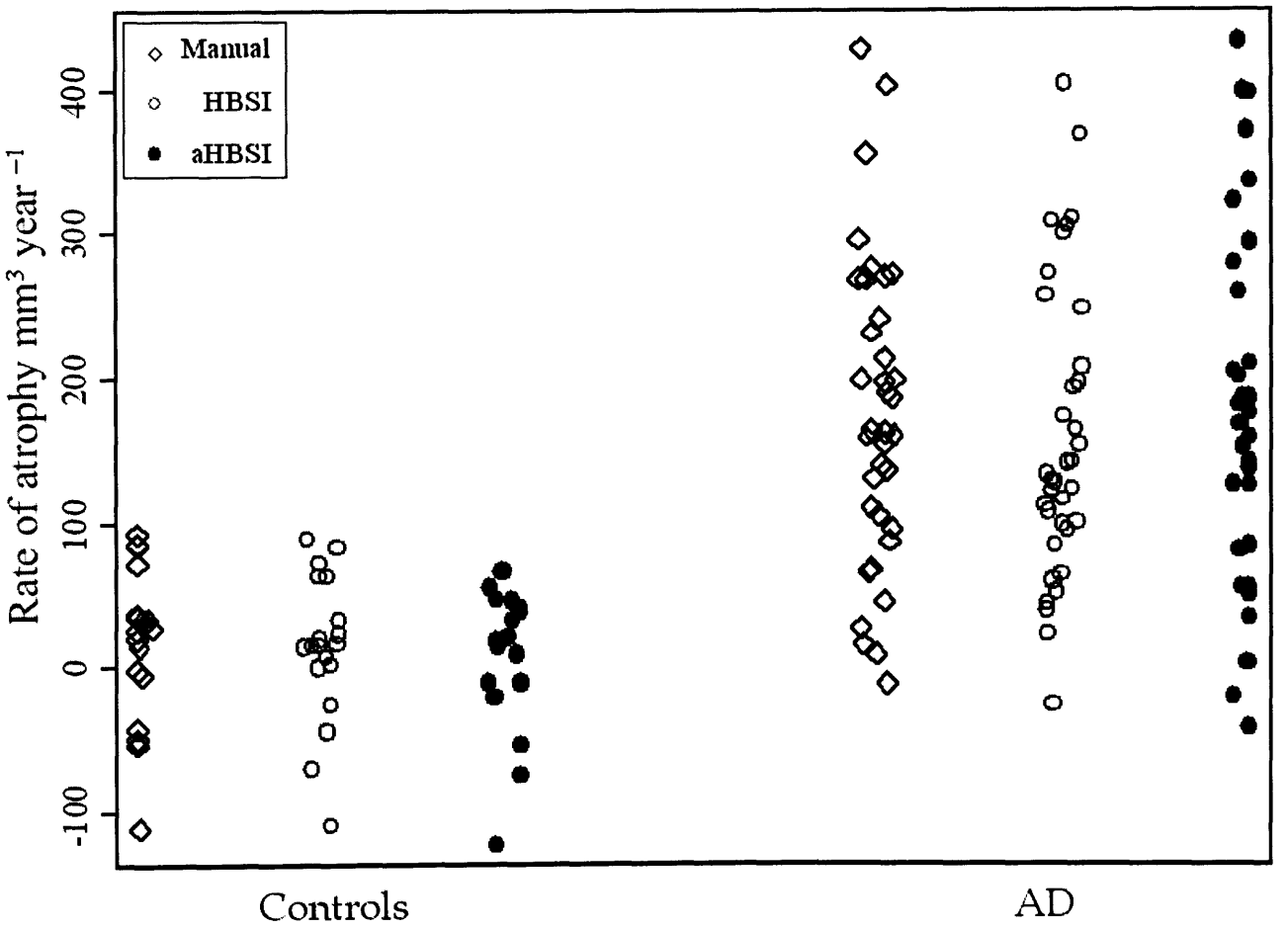


Figure 9.4 Manual, semi-automated HBSI and fully-automated standard aHBSI measures in control and AD groups.

All three methods of generating hippocampal atrophy significantly discriminated AD from controls ($p < 0.01$) (see Figure 9.6). The combination of manual and HBSI was a significantly better discriminator than HBSI alone ($p = 0.011$) and manual alone ($p = 0.023$), suggesting that the two methods together could improve discrimination. The combination of manual and standard aHBSI was significantly better than standard aHBSI alone ($p = 0.0033$) and significantly better than manual alone ($p = 0.01$), suggesting again that the two methods together could improve discrimination. For a 90% specificity cut-off, sensitivities were 72% for manual, 78% for HBSI and 83% for standard aHBSI. For an 80% specificity cut-off, sensitivities were 81% for manual, 86% for HBSI and 83% for standard aHBSI. Power calculations showed that in a trial with 90% power to detect a 20% reduction in atrophy rate of the AD group at the 5% level, 196 subjects per treatment arm would be required of manually-derived measures, compared with 212 for HBSI and 270 for standard aHBSI.

9.3.2. *Influence of differing parameters in the generation of the automated HBSI*

9.3.2.1.Manipulated aHBSI

Table 9.3 shows the initial voxel similarity results for the single-person template manipulations assessed in the subset of the cohort (three AD subjects and three normal controls).

Region 4 (corresponding to the 70% thresholding and two conditional dilations if voxels were between 70-110% mean brain intensity) was the most accurate over these six subjects so this was extended to all baseline segmentations.

Table 9.3 Mean (SD) voxel similarity results for the manipulation of standard single-person template segmentations.

	Control	AD	Total
Single-person template	0.430 (0.077)	0.480 (0.066)	0.455 (0.072)
Region 1	0.524 (0.064)	0.491 (0.025)	0.507 (0.047)
Region 2	0.530 (0.055)	0.481 (0.032)	0.506 (0.048)
Region 3	0.609 (0.063)	0.573 (0.050)	0.591 (0.054)
Region 4	0.622 (0.055)	0.592 (0.065)	0.607 (0.056)

Mean (SD) voxel similarity for the whole group based on this type of manipulation was 0.62 (0.07) for the control group and 0.61 (0.07) for the AD group. The resultant rates of atrophy based on these regions are shown below in Table 9.4 and Figure 9.5. All methods showed significant differences in rates of atrophy between control and AD groups ($p<0.001$) which is demonstrated in Figure 9.6.

There was evidence of a difference between the HBSI and manipulated aHBSI ($p=0.034$) with the SD of the difference from the HBSI being 36.85mm^3. However, there was no difference between the manipulated aHBSI and the manual difference ($p=0.997$) with the SD of the difference being 87.73mm^3.

Table 9.4 Mean (SD) annualised absolute rates of atrophy in controls and AD using manual, semi-automated HBSI, standard automated HBSI (aHBSI) and manipulated aHBSI (mHBSI).

	Manual	HBSI	aHBSI	mHBSI
Controls	18.1 (53.5)	15.3 (50.2)	11.3 (50.4)	15.3 (56.5)
AD	174.6 (106.5)	159.4 (101.2)	172.1 (123.1)	176.0 (114.6)

Owing to the relative accuracy of segmentation (see Figure 9.7, Figure 9.8 and Figure 9.9 for plots of association and Bland-Altman plots of manipulated region volumes against manual volumes in left, right and total hippocampus, and Figure 9.10 for group separation of manual and manipulated baseline volumes), relative rates of atrophy were also calculated using this technique using the manipulated baseline volumes (see Table 9.5). There was no evidence of a difference in the rates of atrophy calculated between manual and manipulated aHBSI methods (p=0.71), with the SD of the difference being 2.16%. The association of these two measures is shown in Figure 9.11.

Table 9.5 Mean annualised relative rates of atrophy in controls and AD using the manual, standard HBSI, and aHBSI calculated from the manipulated baseline region (mHBSI).

	Manual	HBSI	mHBSI
Controls	0.28 (0.90)	0.32 (0.93)	0.28 (1.01)
AD	4.06 (2.61)	4.57 (2.98)	4.43 (3.06)

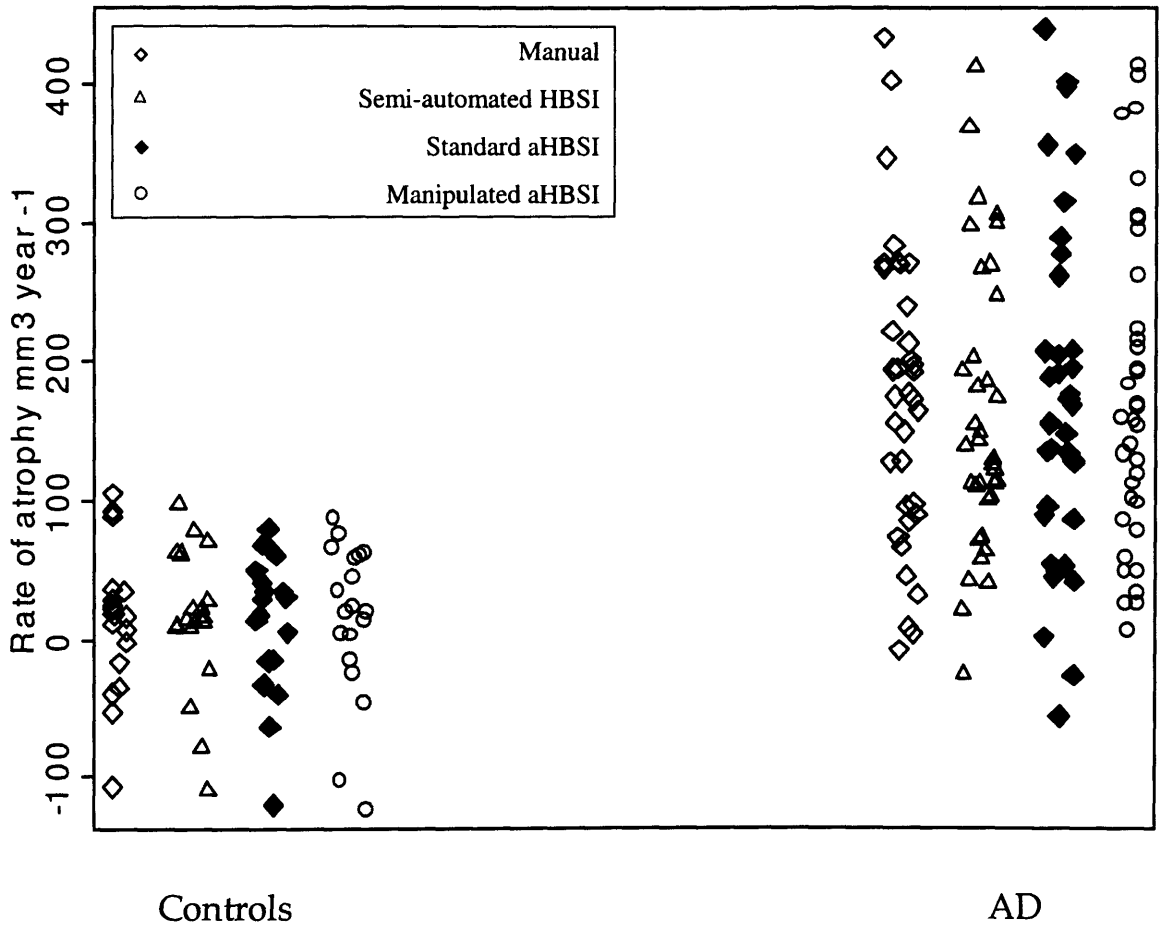


Figure 9.5 Figure demonstrating difference in AD and control groups using manual measures, semi-automated HBSI, standard aHBSI and HBSI generated from manipulating the single-person template automated hippocampal region (manipulated aHBSI).

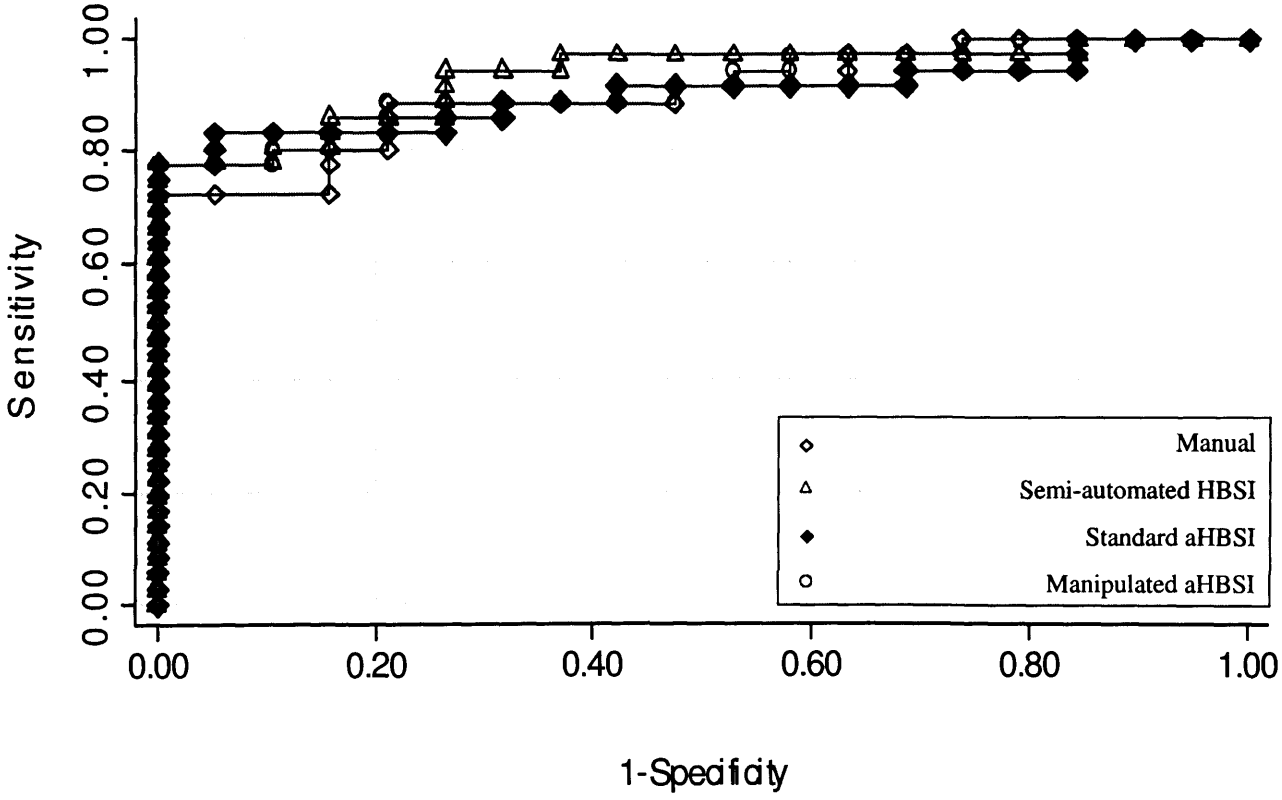


Figure 9.6 Overlaid ROC curves showing AD vs. control group separation using manipulated single-person template aHBSI in comparison to manual, semi-automated HBSI and standard aHBSI.

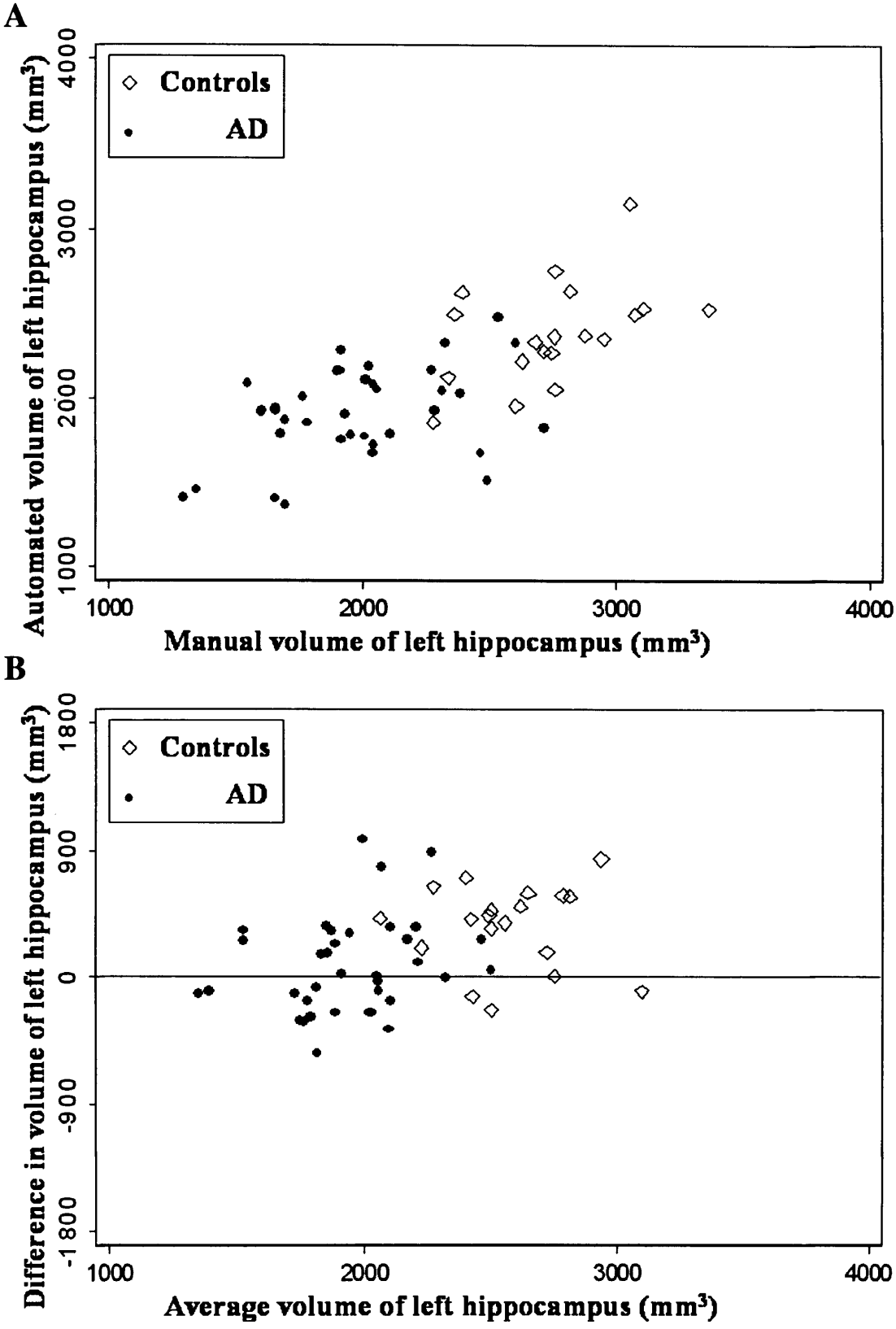


Figure 9.7 A. Association of automated manipulated and manual left hippocampal volumes in AD and control subjects (R^2 for combined subject groups = 0.48). B. Bland-Altman plot of automated manipulated and manual left hippocampal volumes in AD and control subjects.

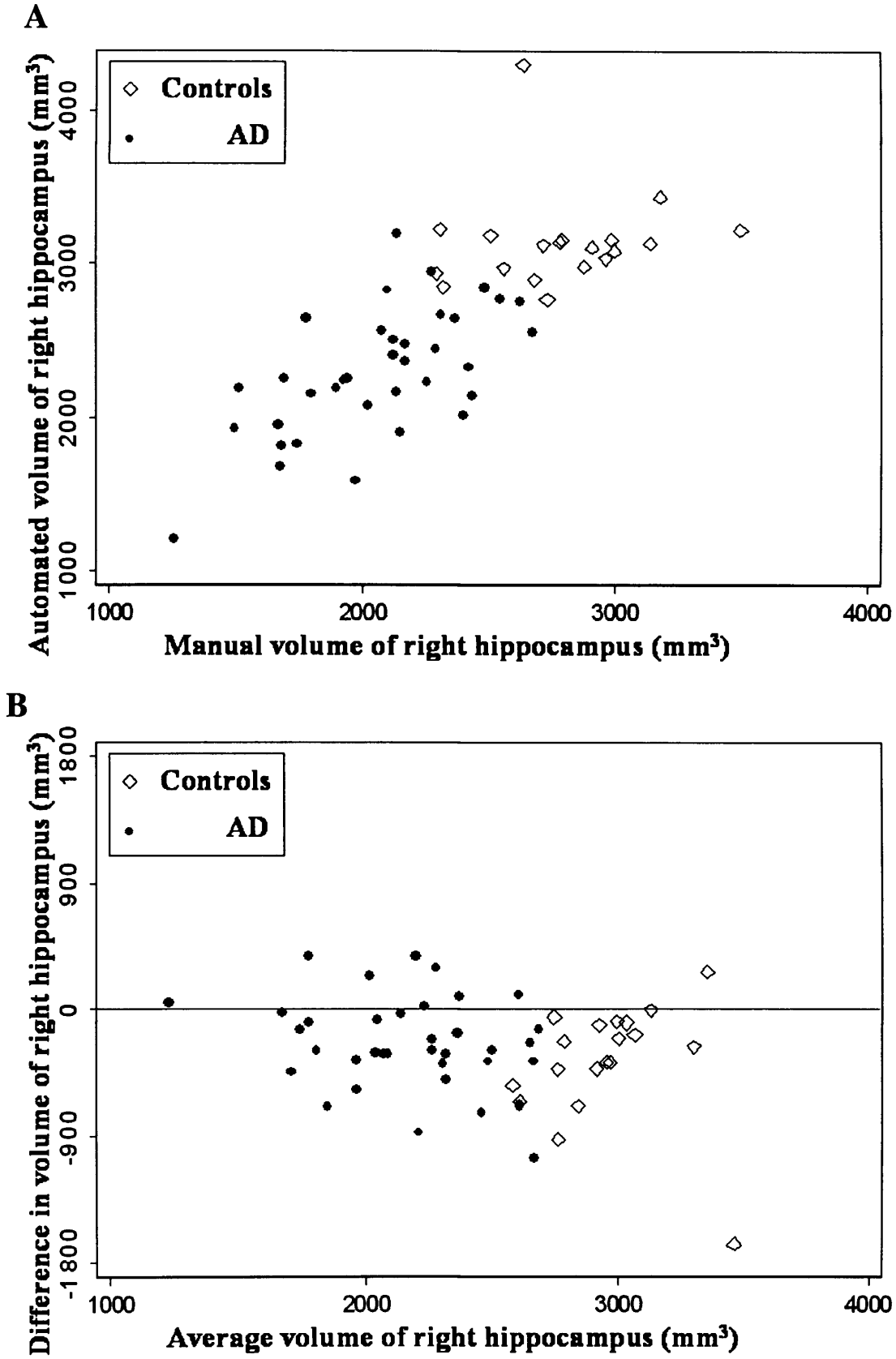


Figure 9.8 A. Association of automated manipulated and manual right hippocampal volumes in AD and control subjects (R^2 for combined subject groups = 0.58). B. Bland-Altman plot of automated manipulated and manual right hippocampal volumes in AD and control subjects.

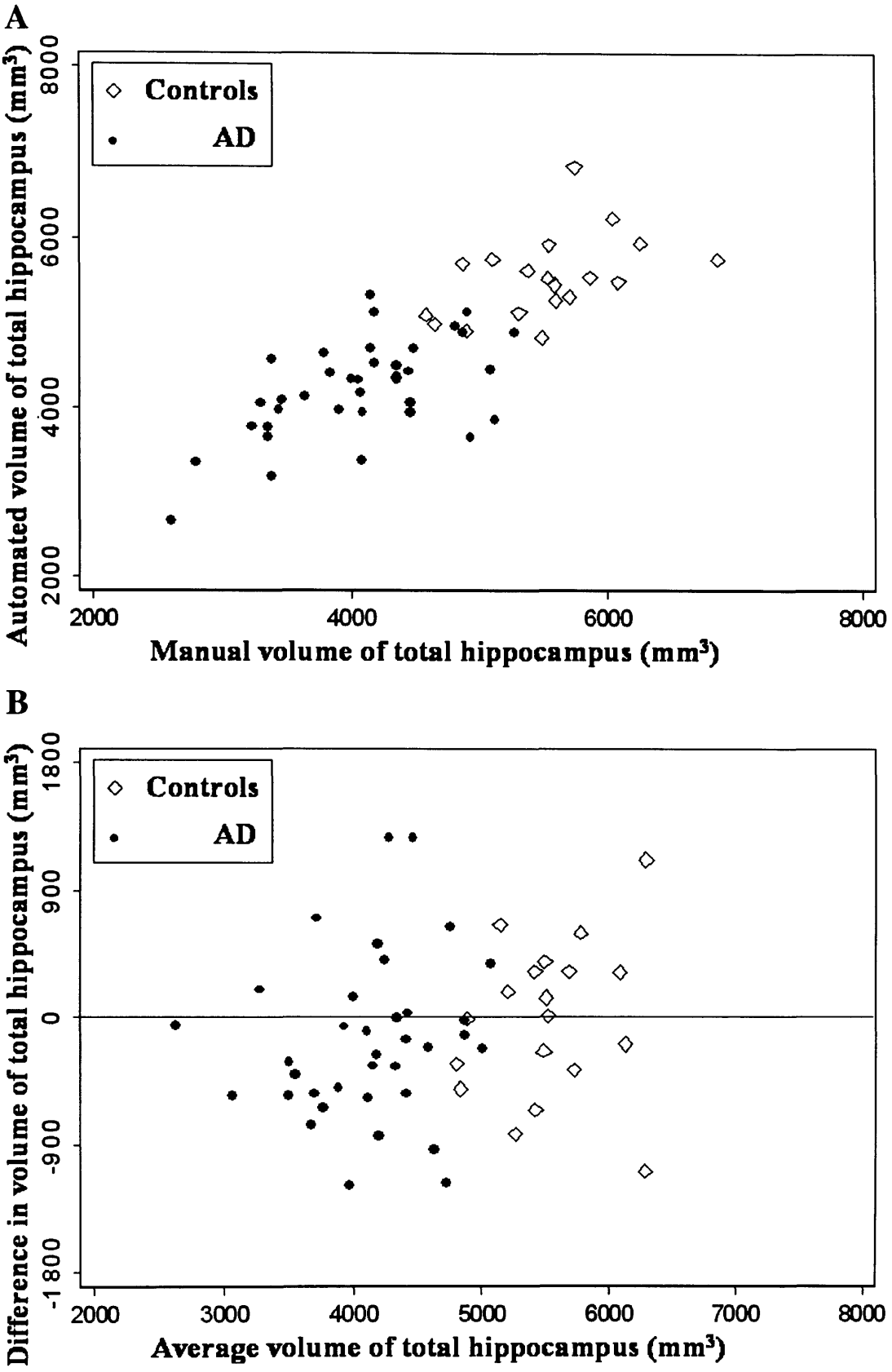


Figure 9.9 A. Association of automated manipulated and manual total (left plus right) hippocampal volumes in AD and control subjects (R^2 for combined subject groups = 0.65). B. Bland-Altman plot of automated manipulated and manual total (left plus right) hippocampal volumes in AD and control subjects.

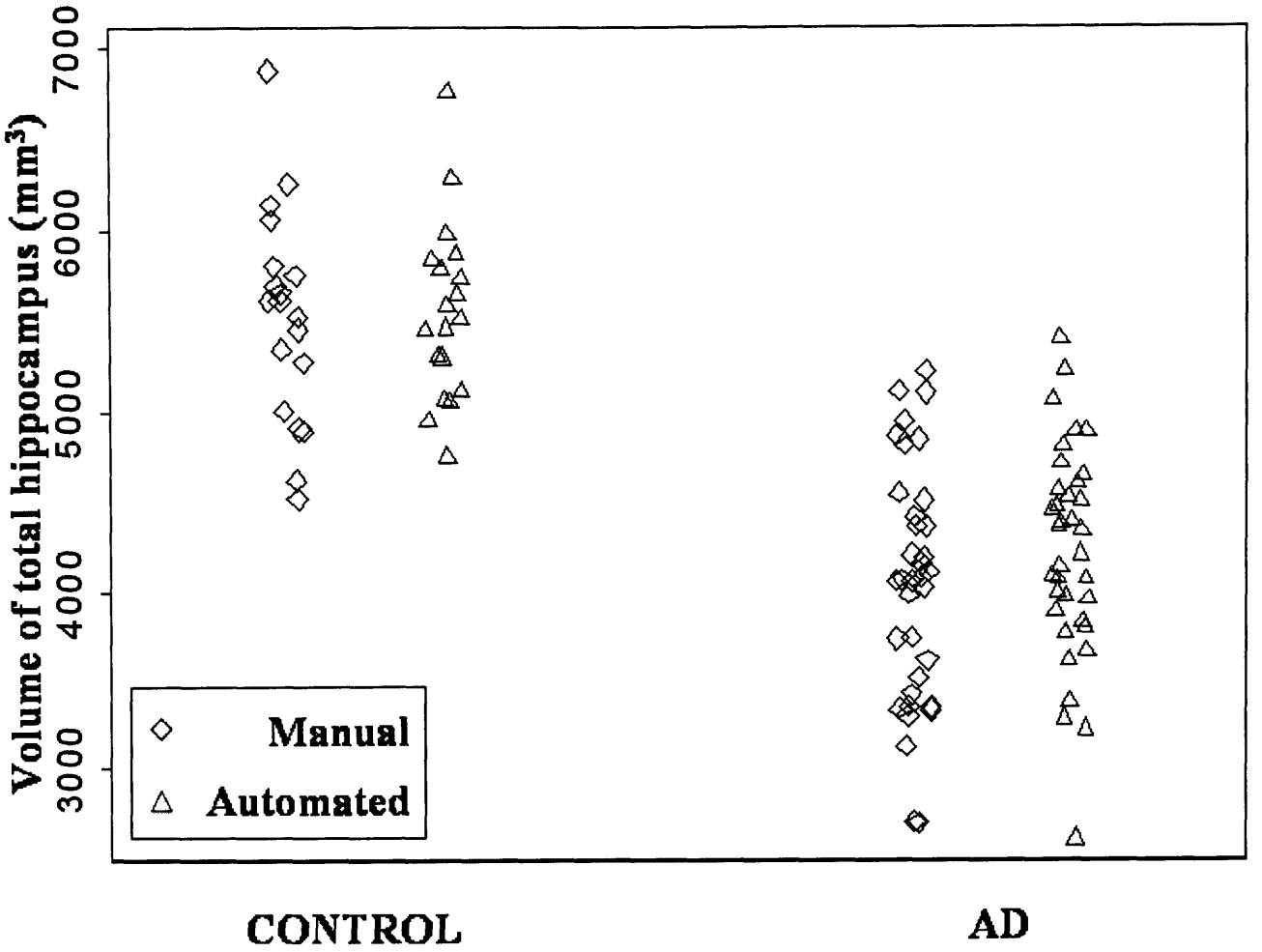


Figure 9.10 Manual and manipulated automated baseline hippocampal volumes in control and AD subjects. Volumes adjusted for TIV

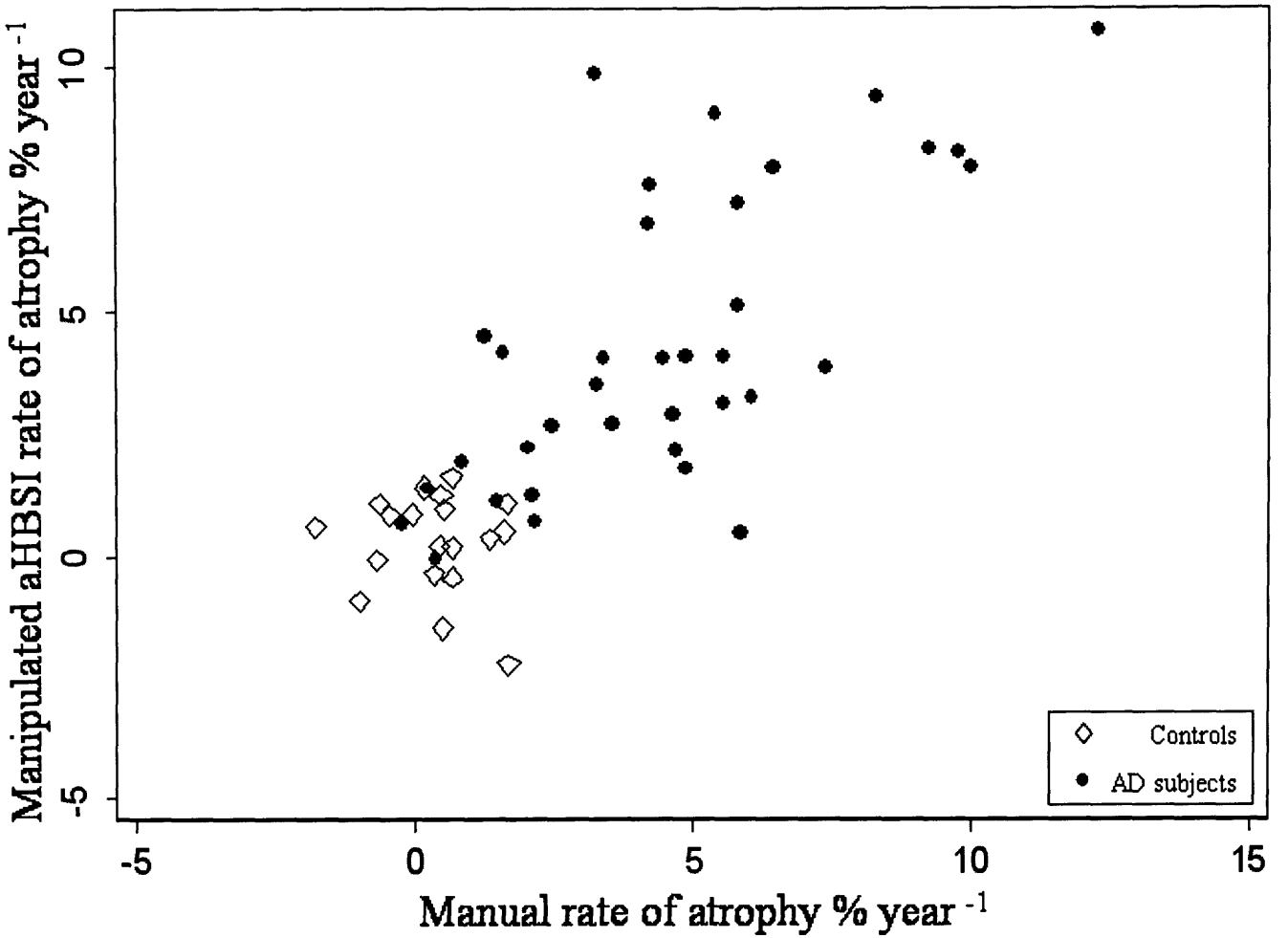


Figure 9.11 Association of automated HBSI (using a single-person template manipulated region) and manual rate of atrophy. R^2 = 0.62 for both AD and control subjects together. The comparative R^2 for the mHBSI and the aHBSI was 0.92.

9.3.2.2.Multi-subject aHBSI.

The accuracies of segmentation of each probability-thresholded region (denoted by the percentage region used followed by aHBSI) are displayed in Table 9.6. Rates for each level of hippocampal probability (denoted by the percentage region used followed by automated HBSI) are shown in Figure 9.12 together with semi-automated HBSI and standard aHBSI and manual measures. Mean (SD) rates are shown below in Table 9.7). All methods showed significant differences in rates of atrophy between control and AD groups ($p<0.001$) which is demonstrated in Figure 9.13. Rates of atrophy differed between the semi-automated and group template region-based HBSI at every probability level ($p<0.03$). The probability region of 60% gave rates which were most similar to the semi-automated HBSI measures with the SD of the difference from the semi-automated being 36.84mm^3.

Table 9.6 Mean (SD) voxel similarity results for the average template region.

	Control	**AD**	**Total**
Standard single-person template	0.50 (0.06)	0.44 (0.44)	0.46 (0.07)
100% aHBSI	0.43 (0.03)	0.29 (0.05)	0.35 (0.08)
80% aHBSI	0.55 (0.05)	0.42 (0.09)	0.47 (0.10)
60% aHBSI	0.50 (0.05)	0.40 (0.11)	0.44 (0.10)

Table 9.7 Mean (SD) absolute rates of atrophy using the multiple-person template hippocampus probability thresholded at differing levels.

	Manual	**HBSI**	**aHBSI**	**100% aHBSI**	**80% aHBSI**	**60% aHBSI**
Controls	20.4 (52.3)	10.3 (62.1)	2.6 (59.7)	14.7 (147.4)	-3.7 (76.1)	-4.4 (50.9)
AD	171.9 (105.2)	145.2 (106.9)	173.9 (111.9)	470.4 (291.9)	264.3 (181.4)	182.6 (116.6)

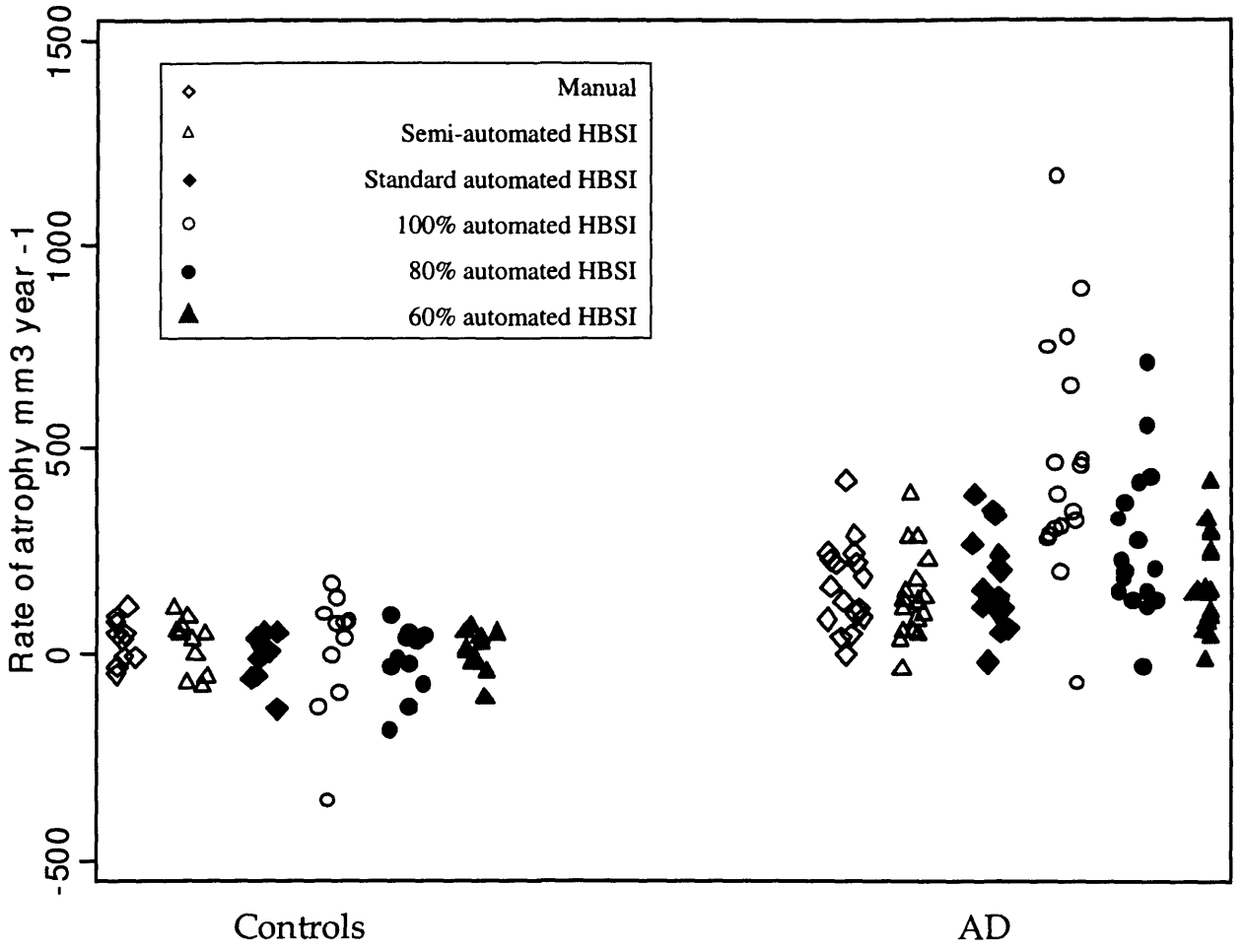


Figure 9.12 Figure demonstrating difference in AD and control groups using manual measures, semi-automated HBSI, standard automated HBSI and HBSI generated from multiple person template regions thresholded at different levels.

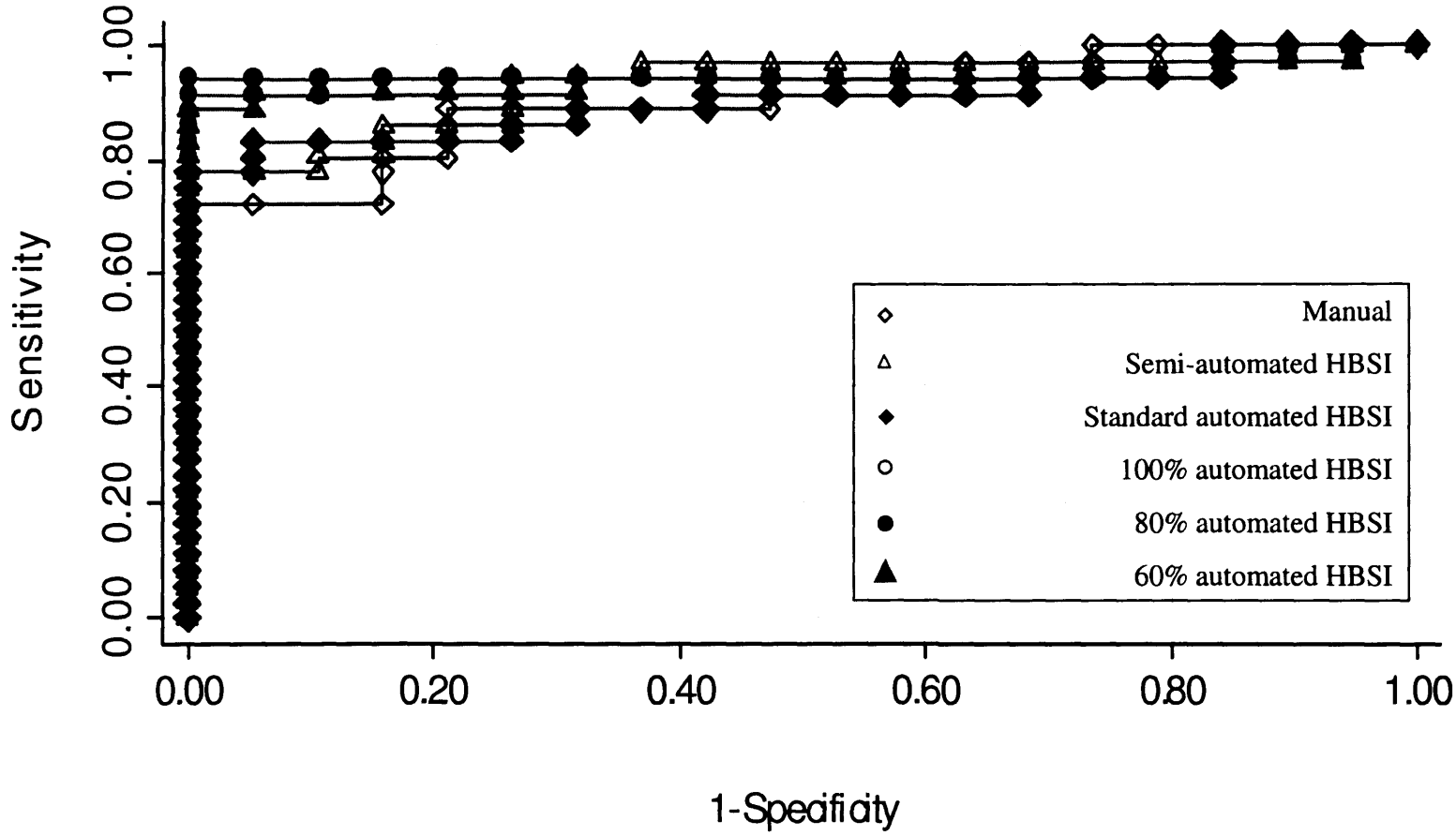


Figure 9.13 Overlaid ROC curves showing AD vs. control group separation using probability thresholded group template regions in comparison to manual, semi-automated HBSI and standard automated HBSI.

9.4. Discussion

In this chapter a way of automating the calculation of hippocampal atrophy rates in AD and control subjects is described. The standard aHBSI method utilised a template hippocampal region and did not require manual segmentation of each subjects' scans. The standard aHBSI method gave similar results to a semi-automated method of generating atrophy rates (HBSI), and to manual measures. There was no significant difference in discriminating the AD and control groups using any method, although it may be that a combination of one of the HBSI methods and manual measures may be superior in separating AD patients from controls compared with any one of the methods alone.

The results presented here for the standard aHBSI are on an absolute scale (mm^3 loss per year). On a relative scale, the findings are in keeping with those in the literature with respect to rates of atrophy in elderly controls, and AD subjects. Mean atrophy rates were around $170mm^3$ per year in AD (which approximates to 4-5% per year) compared with less than $20mm^3$ per year (under 0.5% per year) in controls. This is within the range of previously published hippocampal rates of atrophy for AD patients of this age group (see Table 1.3).

Other techniques exist for automatically or semiautomatically segmenting the hippocampal region. These methods are summarised in Table 2.1. One of these studies has reported atrophy rates by applying their technique to serial imaging (Wang *et al.*, 2003). This technique also used a single-subject template hippocampus and applied non-linear registration methods to propagate this label onto baseline and repeat scans for 18 subjects with mild dementia and 26 controls. This technique requires the application of a number of operator-placed landmarks for the technique to be successful with atrophy rates calculated to be approximately 5% (200mm3) per year in the mildly demented subjects and 2% (100mm3) per year in the normal controls. This rate is similar to that which was found in our AD subjects however the rates reported in controls were lower. This may be a reflection of the different methodologies employed by the two studies or differences in the composition of the control groups, for example, our control group was approximately 5 years younger than that reported by Wang et al. (Wang *et al.*, 2003). Another study used similar techniques to (Wang *et al.*, 2003) and

gave relatively comparable mean atrophy rate results in both AD and control groups (Du *et al.*, 2004).

The influences of a number of factors in the automatic segmentation of the baseline hippocampal regions were also assessed. Manipulating the automatic segmentation of the baseline region using conditional dilations and intensity-thresholding improved the accuracy of the baseline hippocampal regions. This technique also slightly improved the agreement of this HBSI measure with manual segmentations compared with a similar comparison of standard aHBSI and manual measures. The volumes of these manipulated automated regions were accurate enough to use to generate a relative rate of atrophy. Using this technique showed no differences between manual and the automated techniques.

Using an average region generated from half of the subjects did not greatly improve the accuracy of the generated region compared with the manual regions. As a result, this did not improve upon the generation of HBSI measure when comparing with the semi-automated HBSI. In this case, only an average of the hippocampal ROIs were created, rather than the images. It may be that some form of combining variability of both images and regions, such as complex statistical models, may produce templates which can allow more accurate segmentation of the hippocampal region on other scans.

Other features of this template generation and transference of hippocampal labels may affect both the accuracy of segmentation and subsequent HBSI generation. These include the choice of template image, the type of registration method used and the interpolation scheme utilised in the registration process of these images and regions throughout the inter- and intra-subject registration stages. One obvious step to consider is to create a library of manually segmented hippocampi which may be used to segment new scans. By testing the correspondence of the hippocampal region with the library, the most accurate match can be chosen as the single-subject template.

Despite finding large and significant differences in AD subjects' rates of atrophy from controls a large degree of overlap between subject groups was still observed. With respect to control subjects this may reflect true group heterogeneity, scan artefacts or measurement error. In the AD group this may also reflect different disease progression or stage of disease. In this study there was only confirmation of disease in two of the

AD group and one of the controls, therefore there may be some uncertainty as to the pathological status of all of the subjects.

Most of the results presented here could not be shown accurately on a relative scale since the baseline segmentation of the hippocampus using linear inter-subject registration and region transformation did not give accurate volumes. This was demonstrated by the voxel similarity being relatively low for the standard automated region compared with the manual region. This is in keeping with other studies which suggest that affine hippocampal segmentation methods have a similarity measure of around 0.4 for cohort-based atlases (Carmichael *et al.*, 2005; Crum *et al.*, 2004b). The results achieved for the standard aHBSI methodology were marginally better than 0.4 and this may be partly due to the extra hippocampus-hippocampus affine step in the inter-subject stage of the registration procedure. Techniques such as non-linear registration and region propagation may give improved segmentation accuracy of the baseline hippocampi enabling relative rates to be accurately calculated (Carmichael *et al.*, 2005; Crum *et al.*, 2004b). It was found that by manipulating the region using a combination of intensity thresholding and conditional dilations that this segmentation accuracy improved. The level of accuracy achieved is similar to that shown my non-linear registration techniques such as fluid-based registration (Carmichael *et al.*, 2005; Crum *et al.*, 2004b).

One potential problem in the intra-subject stage of the methodology that may affect both HBSI and automated HBSI is the possible mis-registration of baseline and repeat hippocampi. Such potential misalignment of hippocampal borders may cause errors in calculating boundary shifts in such a small area of the brain. However, all intra-subject registrations were visually assessed and this did not appear to be an issue.

In one dataset, the HBSI has previously been shown to be at least as good as manual segmentation in differentiating AD from controls (see Chapter 8, page 204). In this study the automated HBSI and HBSI gave results that were similar to the manual measures. In addition, manual measures were generated by one operator giving this measure a degree of consistency that may not be possible in a multi-centre clinical trial where numerous operators are employed and different scanners will present scans with different intensities and contrasts. It may be that the automated method is most beneficial in a clinical trial situation. Furthermore, this automated hippocampal analysis

in this small study took 20 minutes processor time per hippocampal pair to complete on a 3.4 GHz Intel Xeon server, whereas manual delineation of a hippocampal pair took 90 minutes of trained operator time following a period of three months operator training.

9.5. Chapter Conclusion

In conclusion, hippocampal atrophy rates may be calculated using a combination of inter- and intra-subject registration and application of the hippocampal boundary shift integral. This may be of use both diagnostically, and in large studies where many serial scans require analysis.

10. DETECTING THE LOCATION OF HIPPOCAMPAL LOSS IN AD

10.1. Introduction

As discussed in Chapters 1, 4 and 6 differences in volume of the hippocampus in AD compared with controls are frequently reported in the literature, with AD patients having lower hippocampal volumes compared with matched controls (see Table 1.1). Changes in these volumes over time, often reported as a rate of atrophy, are also widely reported, with AD patients having greater changes than controls (see Chapters 1, 5, and 6 and specifically Table 1.3).

The "gold-standard" method of measuring hippocampal volume loss (as used to generate results in Chapters 4, 5 and 6) utilises separate manual measurements at each time-point with the generation of binary masks (where a voxel is either labelled "in" or "out") which are used to calculate volumes. Rate of change in volume over time (atrophy rate) is usually calculated by taking the second volume from the first and normalising this difference to the inter-scan interval. However, summarising the data in this manner loses information about the spatial distribution of atrophy which may have occurred as a result of the disease or its progression, such as which parts of the hippocampus may be preferentially affected by the disease. This information may be useful for early diagnosis, differential diagnosis or to understand the natural history of the effects of the disease.

A number of methods have been utilised for whole-brain analysis to assess where differences in the amount of grey matter can be found when comparing a group of AD subject scans with a group of control scans. One such technique, VBM, has been shown to highlight hippocampal involvement in AD both cross-sectionally (Baron *et al.*, 2001; Karas *et al.*, 2003) and longitudinally (Scahill *et al.*, 2002). Using another group-level method which statistically assesses grey matter volume following cortical pattern matching, it was also found that the temporal lobe including hippocampus was affected by AD (Ballmaier *et al.*, 2004).

One potential problem with using group-level whole-brain techniques is that they require an alignment step which spatially matches the brain regions. As a result, there is evidence that small structures such as the hippocampus may not be accurately aligned (Mosconi *et al.*, 2005). This impacts upon the ability of these whole-brain based

analyses to give reliable results regarding 1) the involvement of the hippocampus in a specific group of scans compared with another or 2) which sections of the hippocampus may be affected. This problem with alignment combined with variability of hippocampal sizes and shapes seen in a disease population may explain how some studies have shown little hippocampal involvement in severe AD (Scahill *et al.*, 2002).

Some work has attempted to circumvent this by spatially matching hippocampi. Thompson's group has published methods of spatially matching hippocampi using a process of whole-brain registration to a template (ICBM152), manual delineation of hippocampi, and surface-based registration of the smoothed regions (Thompson *et al.*, 2004a) (erratum in press). These methods have been shown to detect cross-sectional differences at baseline and repeat imaging in the hippocampi of both AD (n= 17) and control subjects (n=14, mean age approximately 70 years). These differences were located particularly in the left hippocampal head, however, surprisingly, rates of atrophy in the left hippocampus were similar in controls and AD subjects.

Methods have also been published methods for assessing spatial differences in hippocampi using surface deformations which can be applied both cross-sectionally (inter-subject) (Csernansky *et al.*, 2000) and longitudinally (intra-subject and inter-subject) (Wang *et al.*, 2003). These methods involve using a labelled single-person cohort template to propagate hippocampal regions onto the remainder of the cohort whilst storing the deformations required to map the template hippocampus to the cohort. These stored deformations are then analysed using methods similar to principle component analysis which produces maps of average inward or outward deformation of the hippocampal surface. These methods have shown that the hippocampus is deformed in dementia of the Alzheimer type compared with controls and the authors claim that this inward deformation is mainly specific to the hippocampal head and the CA1 region of the hippocampus in AD subjects in their mid-seventies (Csernansky *et al.*, 2000).

These methods of hippocampal matching may be preferable to matching whole-brains assuming the hippocampi are adequately aligned. However, there is still an assumption that remains with all these techniques that there is correspondence between specific parts of the structure in all hippocampi.

The aim of this chapter is to investigate the distribution of atrophy within the hippocampus in AD subjects compared with controls both cross-sectionally and longitudinally.

10.2. Methods

10.2.1. Subjects

Subjects were used from the MIRIAD cohort comprising a group of 36 elderly sporadic AD subjects and 19 age-matched controls. This is the same group of subjects and scans as used in Chapter 9. For subject demographics see Table 9.1 and for further details see Chapter 9 (page 244).

10.2.2. Scan acquisition

T1-weighted volumetric MR brain scans were performed using the MIRIAD protocol (see Appendix Three).

10.2.3. Hippocampal segmentation

All segmentations were performed using MIDAS software. Brain regions on the baseline and repeat images were segmented using a previously described method (Freeborough *et al.*, 1997)(see Appendix Four), and all images were registered into MNI 305 atlas space (Mazziotta *et al.*, 1995) using six degrees of freedom (dof) to ensure all images were in a similar orientation. Following this the repeat images (second time-point) were registered to the baseline images using nine dof registration, to align the brains within a scan-pair accurately. Hippocampi were delineated for each subject on the baseline and nine dof brain-brain registered-repeat images. Mirror-image volumetry (see page 101) was used for the hippocampal segmentation process. The baseline and registered-repeat images were loaded simultaneously with the scans in a random chronological order. The hippocampus was always measured on the right-hand side of the presented image with the investigator blinded to the subject's name, diagnosis, chronological order, and left-right orientation of the scans. The protocol used is detailed in Appendix Four.

10.2.4. Generation of average AD and control maps

All hippocampi were affine registered (12 dof) brain to brain to the individual subject used as the single person template in Chapter 9 (see page 246, and Figure 10.1). The hippocampi were then rigidly registered (six dof) to this labelled template's

hippocampus and the resulting binary regions in template space were stored. The volume over which the cost function for this registration was calculated was the template hippocampal region dilated by two voxels. The cost function used for all registrations was the SD of baseline image / repeat image. One year interval hippocampi were also used transformed into this template space by applying the transformation parameters calculated from transforming the baseline hippocampi into the template space. These masks were again stored. Within each subject group average maps were created from these masks, for left and right hippocampi at baseline and repeat imaging, giving a total of eight hippocampal maps. These were used to overlay results from the voxel-wise significance tests. For purposes of presentation, difference images showing one map minus the other (AD – controls) are also presented.

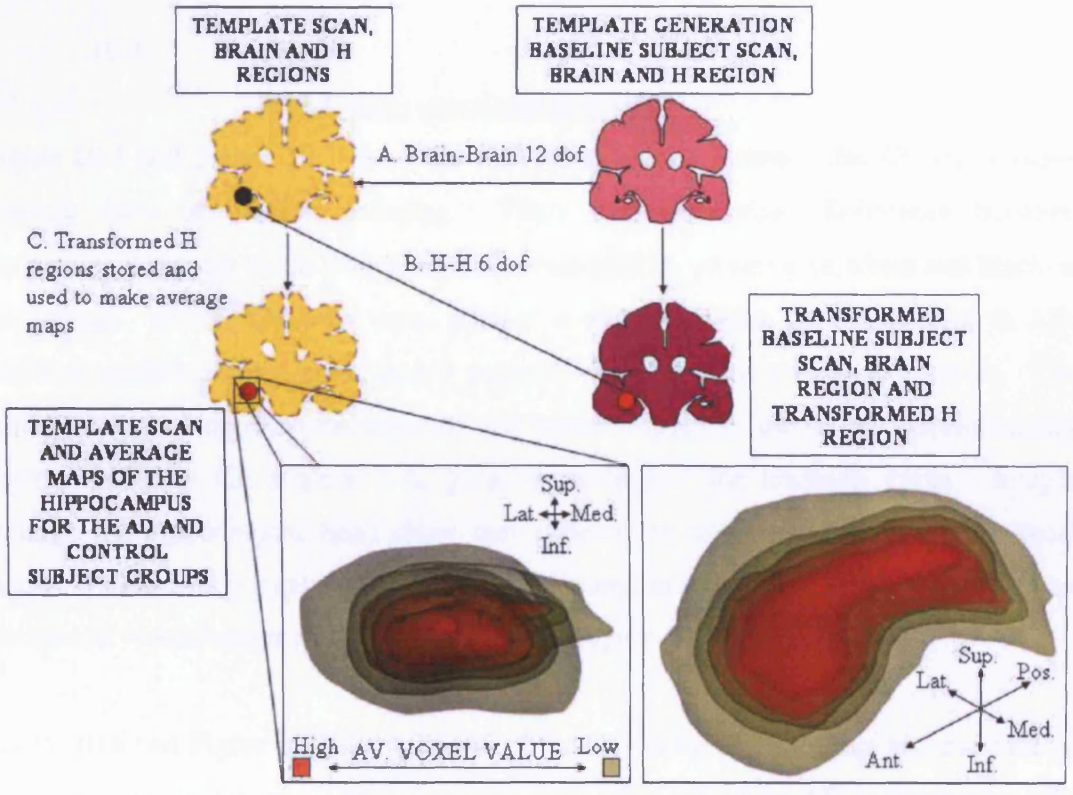


Figure 10.1 Schematic diagram showing creation of the average maps of the hippocampus in AD and control subjects.

H = hippocampus, Sup.= superior, Lat. = lateral, Med.= medial, Inf. = inferior, Pos. = posterior, Ant = anterior.

10.2.5. Statistical analysis

The hippocampal masks transformed into template space were tested both cross-sectionally between subject groups and longitudinally within subject group. Owing to the binary nature of the data, Fisher's exact test was performed at each voxel to establish whether there were cross-sectional differences between subject groups. Results were reported for two-tailed tests as well as left and right tailed tests. This was performed in the left and right hippocampus separately, and at both baseline and repeat scanning time-point. To assess whether changes could be detected over time McNemar's test was used in a similar way, testing voxel by voxel over the mask data sets with subject group over time. This test was performed in both left and right hippocampi. Results of these statistical tests are presented as colour overlays on the average control maps created above.

10.3. Results

10.3.1. Cross-sectional analysis

Figure 10.2 and Figure 10.3 show the difference images between the AD and control average maps at baseline imaging. There are widespread differences between hippocampal regions in the two groups demonstrated by presence of white and black in the images. White represents areas present in controls which are not present in AD. Black represents areas present in AD subjects which are not present in controls. The white areas are located in the superior and lateral regions of the image, approximately corresponding to CA regions 1-4, gyrus dentatus and the uncinate gyrus. Images through the hippocampal head show that differences can also be seen in the uncal sulcus. Areas which appear dark are mainly located in the subiculum. Similar patterns are seen at repeat imaging (see Figure 10.4 and Figure 10.5).

Figure 10.6 and Figure 10.7 show results of the two-tailed Fishers exact test overlaid on the average control maps. These show two main areas where the AD and control masks differ (blue on overlay), in the superior region and a region in the inferior of the hippocampus. The regions which correspond to these differences are those highlighted from the difference images above, namely CA regions 1-4, gyrus dentatus and uncinate gyrus and subiculum. Areas in the uncal sulcus did not appear significantly different. This is possibly due to the variability seen in this region in AD.

Figure 10.8 and Figure 10.9 show the overlay of the left-tailed Fisher's exact test. These blue regions show where AD subjects have significantly fewer voxels labelled hippocampus (blue on overlay) overlaid on the average control maps. This shows that control hippocampi are larger than AD in the superior, medial and lateral regions corresponding to the CA 1-4 regions, the gyrus dentatus, and some areas of the uncinate gyrus.

Figure 10.10 and Figure 10.11 show the overlay of the right-tailed Fisher's exact test. These blue regions show where AD subjects have significantly more voxels labelled hippocampus (blue on overlay) overlaid on the average control maps. This shows that AD hippocampi appear to be larger than controls in inferiorly.

Figure 10.12, Figure 10.13 and Figure 10.14 show two-tailed, left-tailed and right-tailed results respectively in the hippocampi at repeat imaging (blue on overlay) overlaid on the average control maps. Only sagittal images are shown as the areas of significance are very similar to those of the baseline imaging.

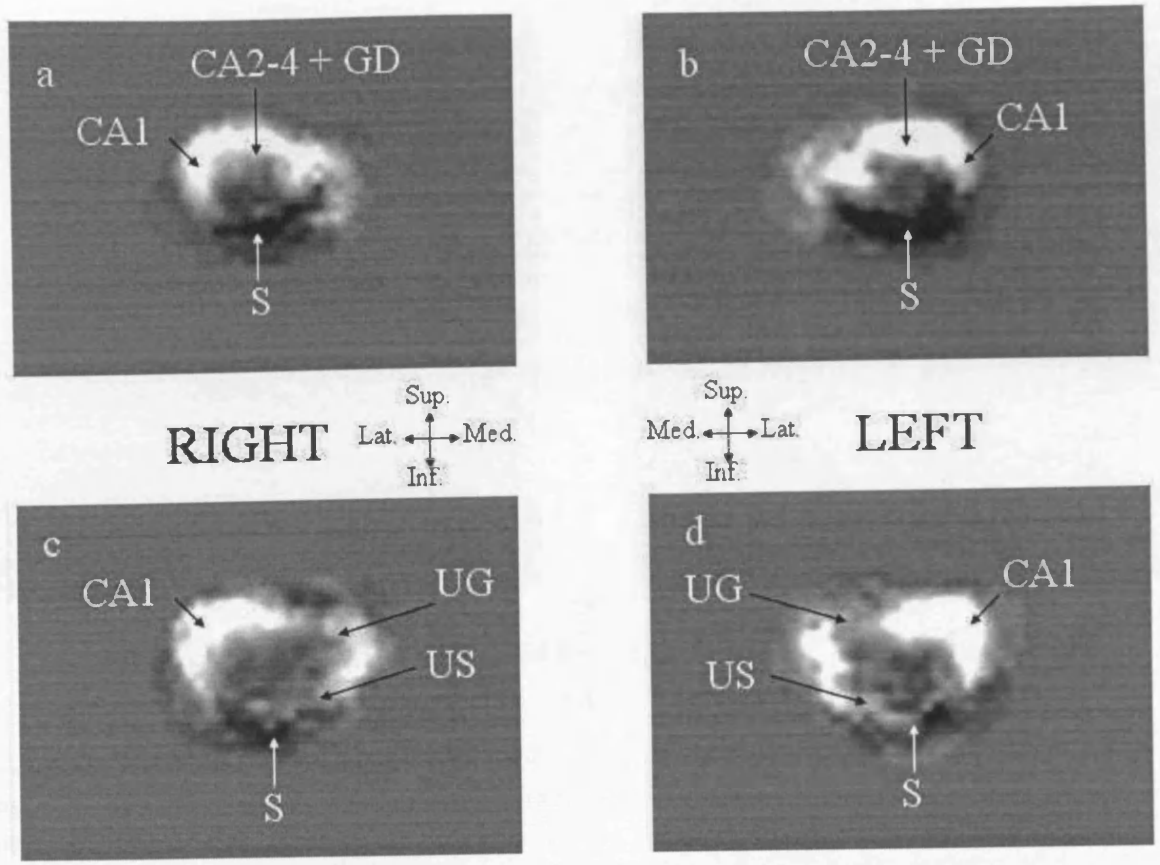


Figure 10.2 Coronal difference images of hippocampal maps of AD vs. controls at baseline imaging.
Images taken through hippocampal body (a and b) and hippocampal head (c and d) in the right (a and c) and left (b and d) hippocampus. White shows where AD < controls and black where AD > controls, grey shows no difference between the two.
Anatomical labels: CA1: cornu Ammonis section 1, CA2-4: cornu Ammonis sections 2-4, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus.

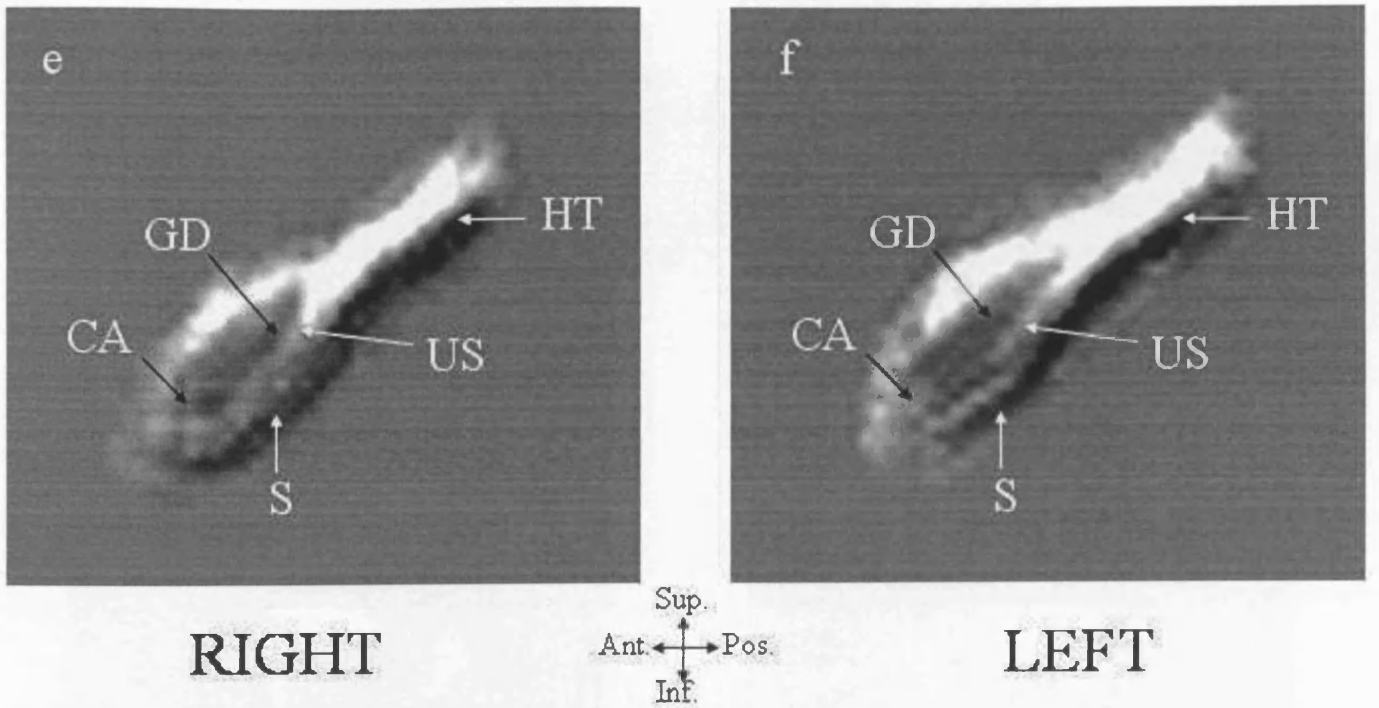


Figure 10.3 Sagittal difference images of hippocampal maps of AD vs. controls at baseline imaging.
Images taken through the hippocampus in the right (e) and left (f) hippocampus. White shows where AD < controls and black where AD > controls, grey shows no difference between the two.
Anatomical labels: CA1: cornu Ammonis section 1, CA2-4: cornu Ammonis sections 2-4, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus.
These imply that in the sagittal view the hippocampus is atrophying along its length and being displaced postero-inferiorly, the overlays also suggest the opening of the uncal sulcus is a characteristic feature of a characteristic feature of the loss.

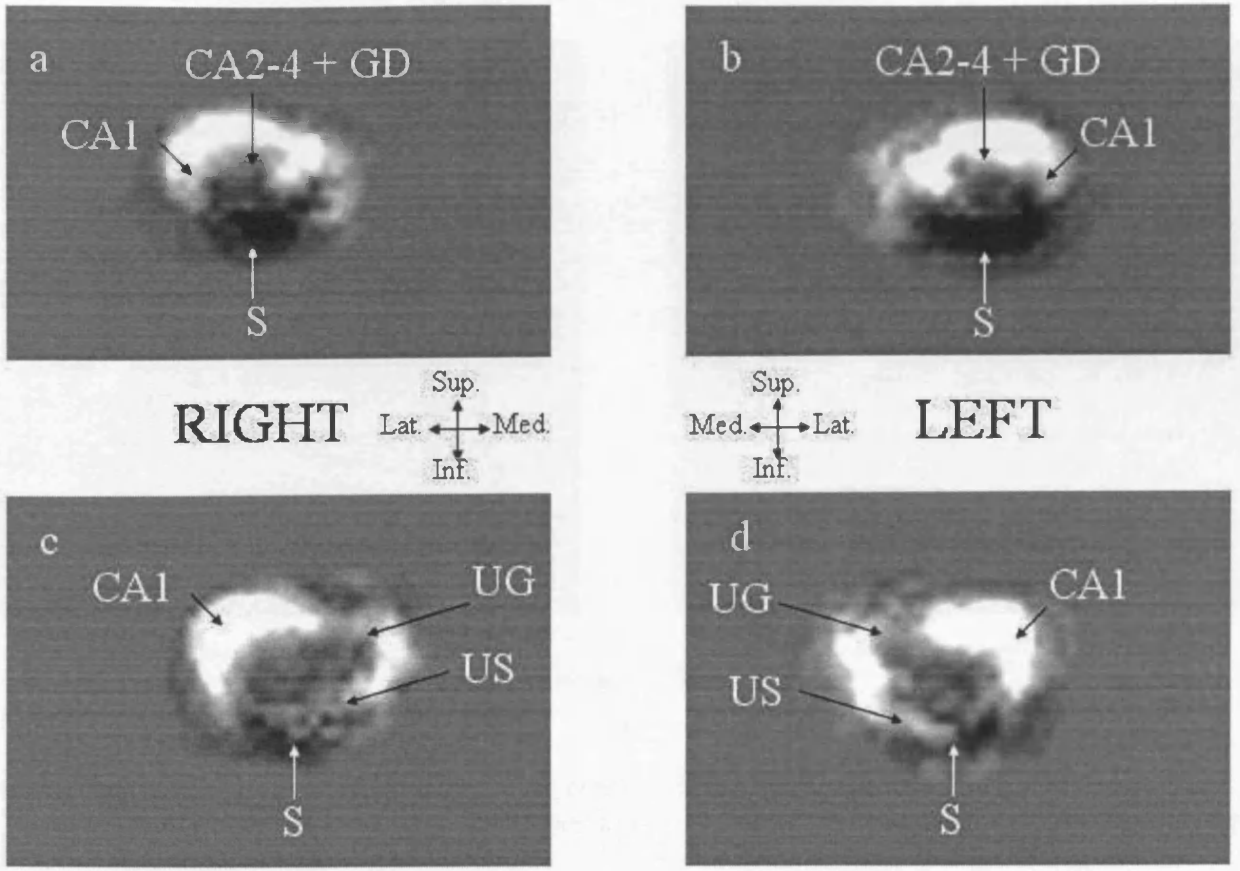


Figure 10.4 Coronal difference images of hippocampal maps of AD vs. controls at repeat imaging.
Images taken through hippocampal body (a and b) and hippocampal head (c and d) in the right (a and c) and left (b and d) hippocampus. White shows where AD < controls and black where AD > controls, grey shows no difference between the two.
Anatomical labels: CA1: cornu Ammonis section 1, CA2-4: cornu Ammonis sections 2-4, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus.

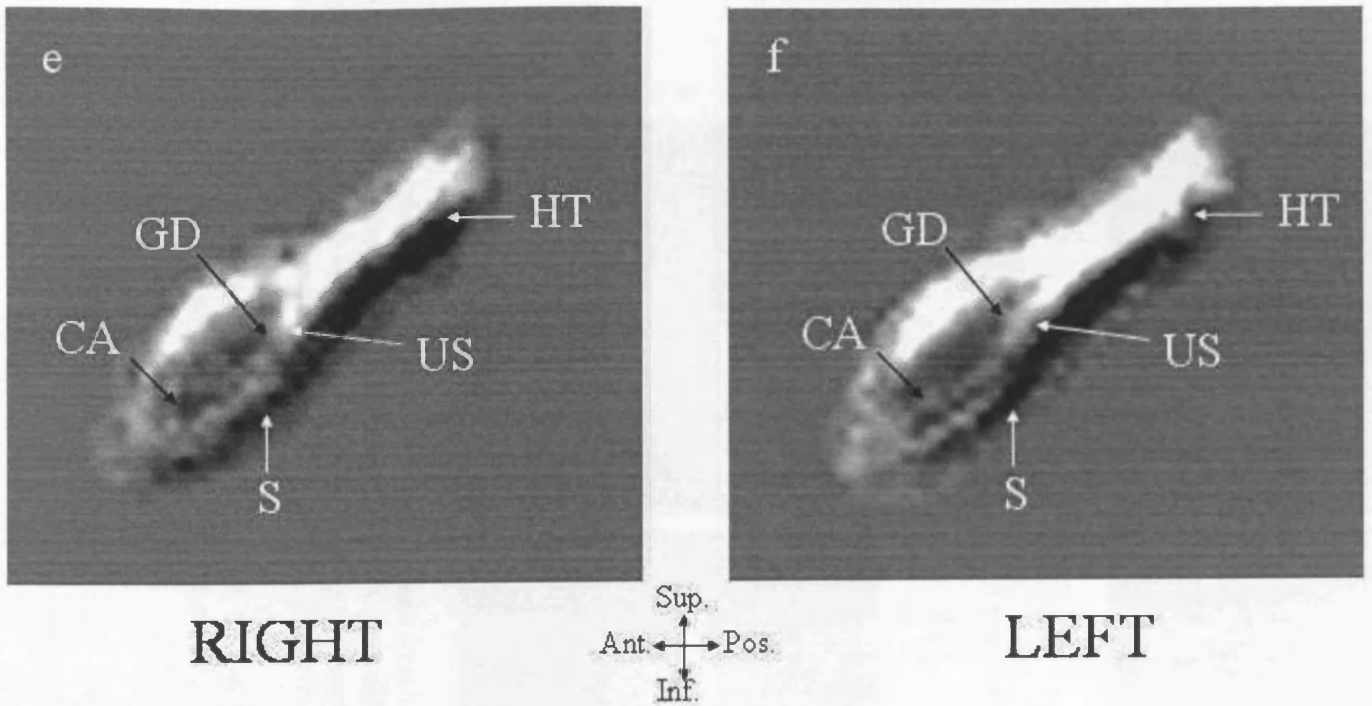


Figure 10.5 Sagittal difference images of hippocampal maps of AD vs. controls at repeat imaging.
Images taken through the hippocampus in the right (e) and left (f) hippocampus. White shows where AD < controls and black where AD > controls, grey shows no difference between the two.
Anatomical labels: CA1: cornu Ammonis section 1, CA2-4: cornu Ammonis sections 2-4, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus.

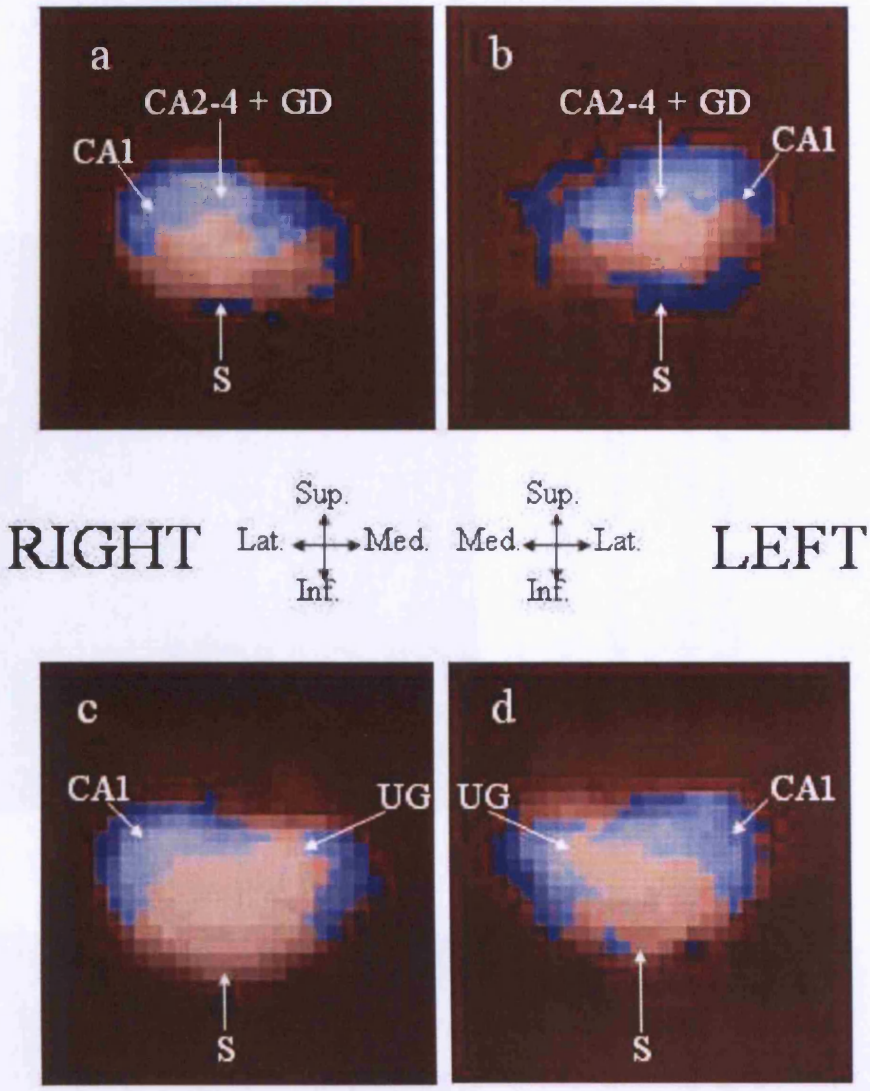


Figure 10.6 Coronal images of control average maps overlaid with Fisher's exact test results showing where AD and controls are significantly different at baseline imaging. Images taken through hippocampal body (a and b) and hippocampal head (c and d) in the right (a and c) and left (b and d) hippocampus. The overlay shows $p \leq 0.05$ where blue and $p > 0.05$ where red.
Anatomical labels: CA1: cornu Ammonis section 1, CA2-4: cornu Ammonis sections 2-4, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus.

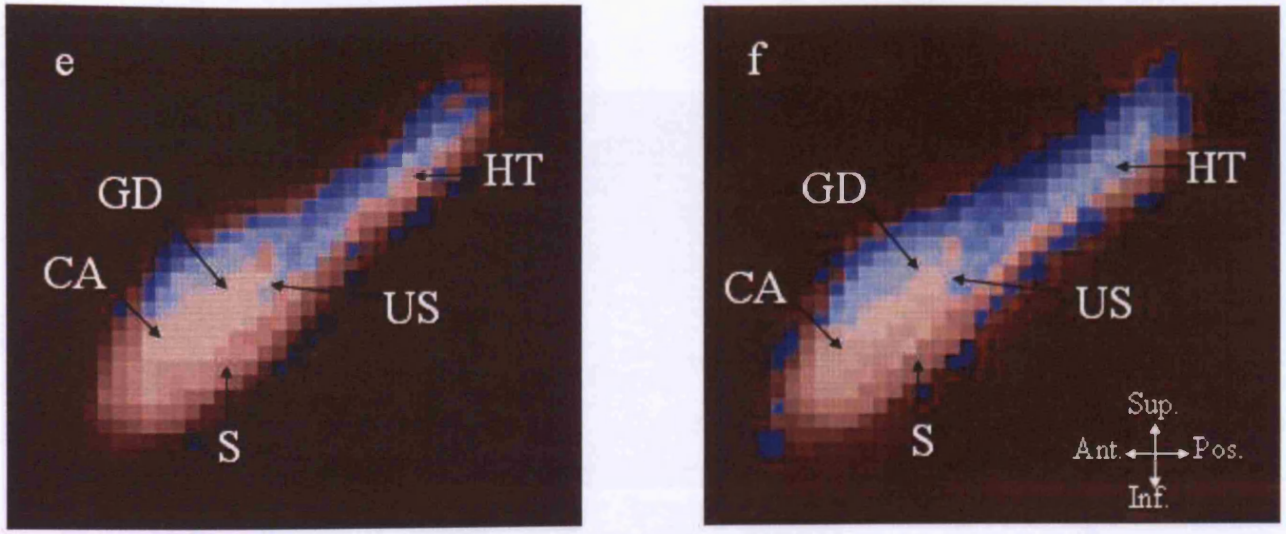


Figure 10.7 Sagittal images of control average maps overlaid with Fisher's exact test results showing where AD and controls are significantly different at baseline imaging.
Images taken of both the right (e) and left (f) hippocampus. The overlay shows $p \leq 0.05$ where blue and $p > 0.05$ where red.
Anatomical labels: CA1: cornu Ammonis section 1, CA2-4: cornu Ammonis sections 2-4, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus.

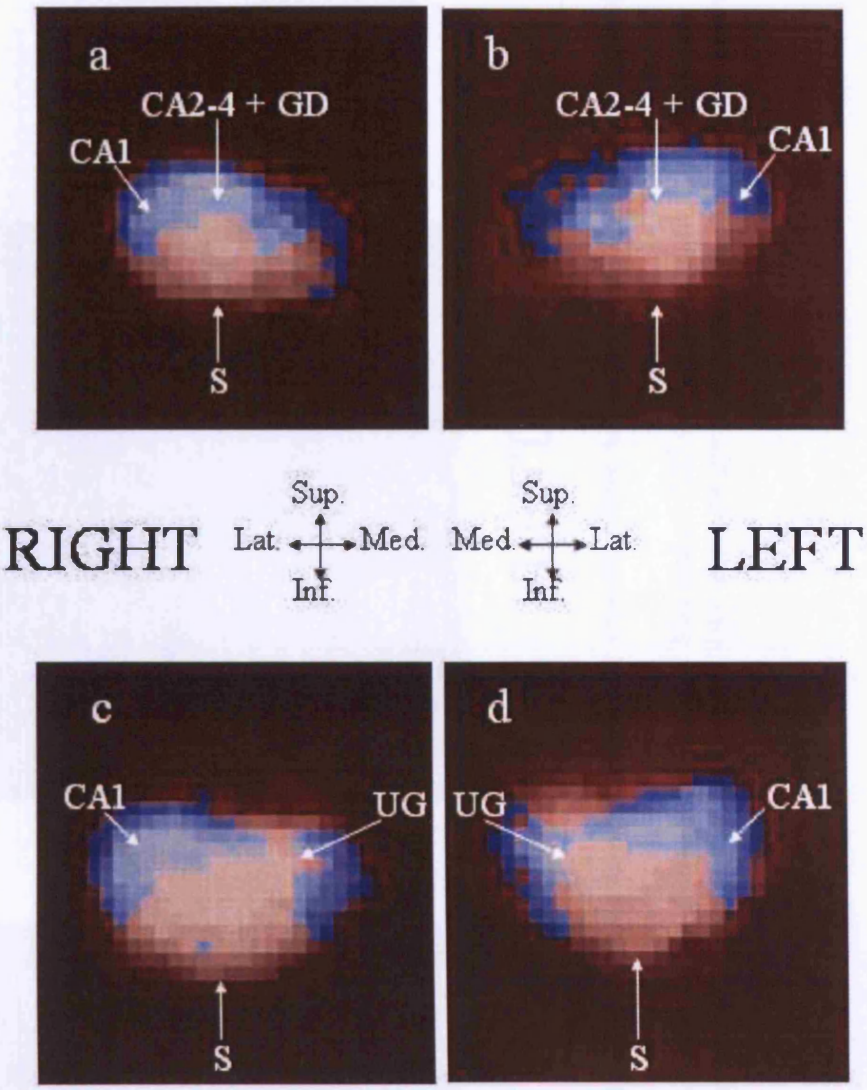


Figure 10.8 Coronal images of control average maps overlaid with Fisher's exact test results showing where AD significantly < controls at baseline imaging.
Images taken through hippocampal body (a and b) and hippocampal head (c and d) in the right (a and c) and left (b and d) hippocampus. The overlay shows $p \leq 0.05$ where blue and $p > 0.05$ where red.
Anatomical labels: CA1: cornu Ammonis section 1, CA2-4: cornu Ammonis sections 2-4, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus.

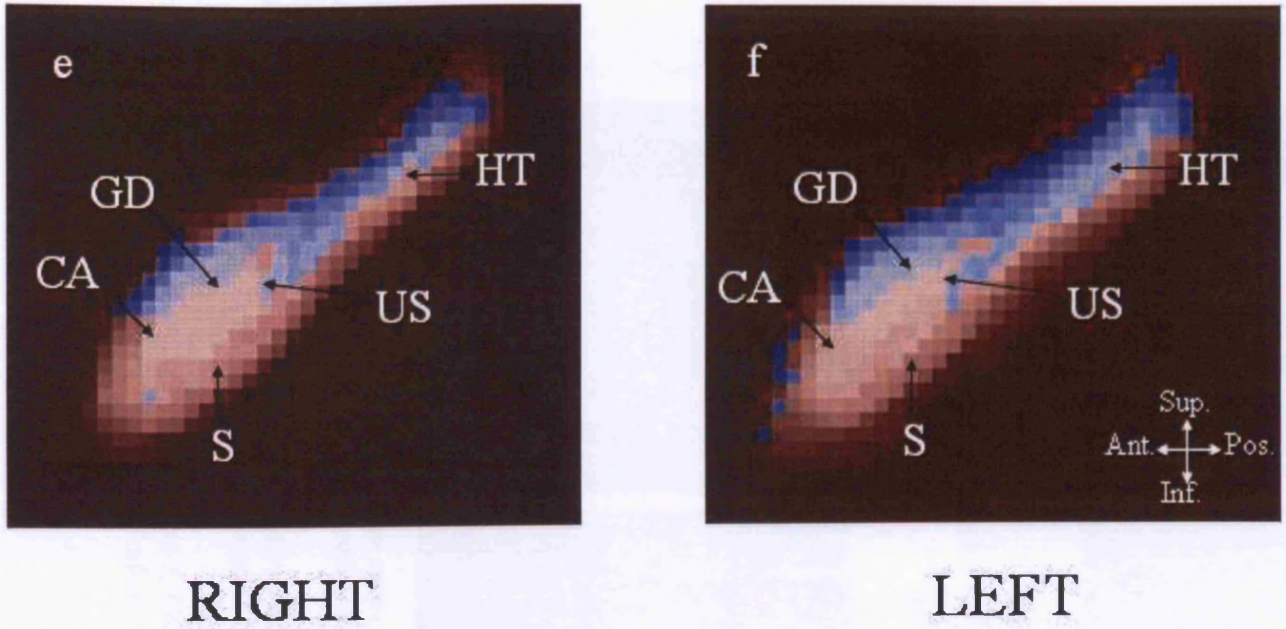


Figure 10.9 Sagittal images of control average maps overlaid with Fisher's exact test results showing where AD significantly < controls at baseline imaging.
Images taken of both the right (e) and left (f) hippocampus. The overlay shows $p \leq 0.05$ where blue and $p > 0.05$ where red.
Anatomical labels: CA: cornu Ammonis, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus, HT: hippocampal tail.

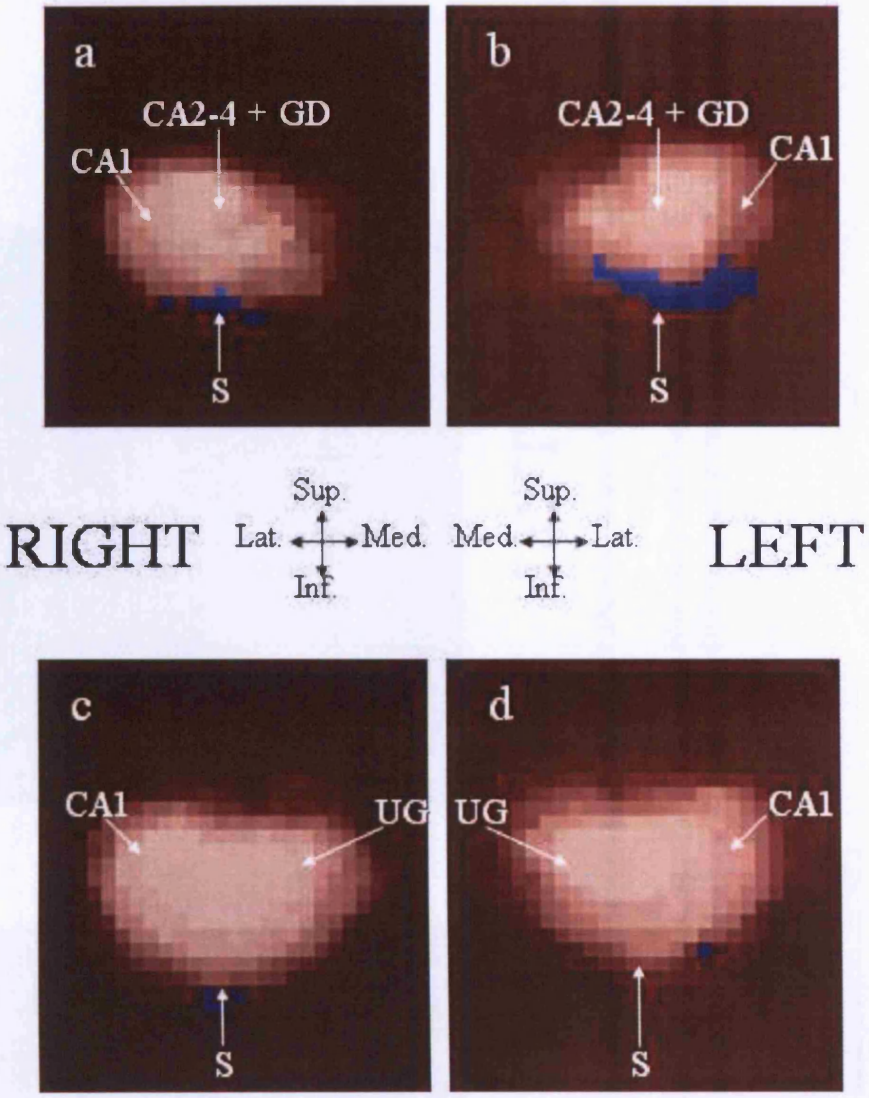


Figure 10.10 Coronal images of control average maps overlaid with Fisher's exact test results showing where AD significantly > controls are significantly different at baseline imaging.
Images taken through hippocampal body (a and b) and hippocampal head (c and d) in the right (a and c) and left (b and d) hippocampus. The overlay shows $p \leq 0.05$ where blue and $p > 0.05$ where red.
Anatomical labels: CA1: cornu Ammonis section 1, CA2-4: cornu Ammonis sections 2-4, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus.

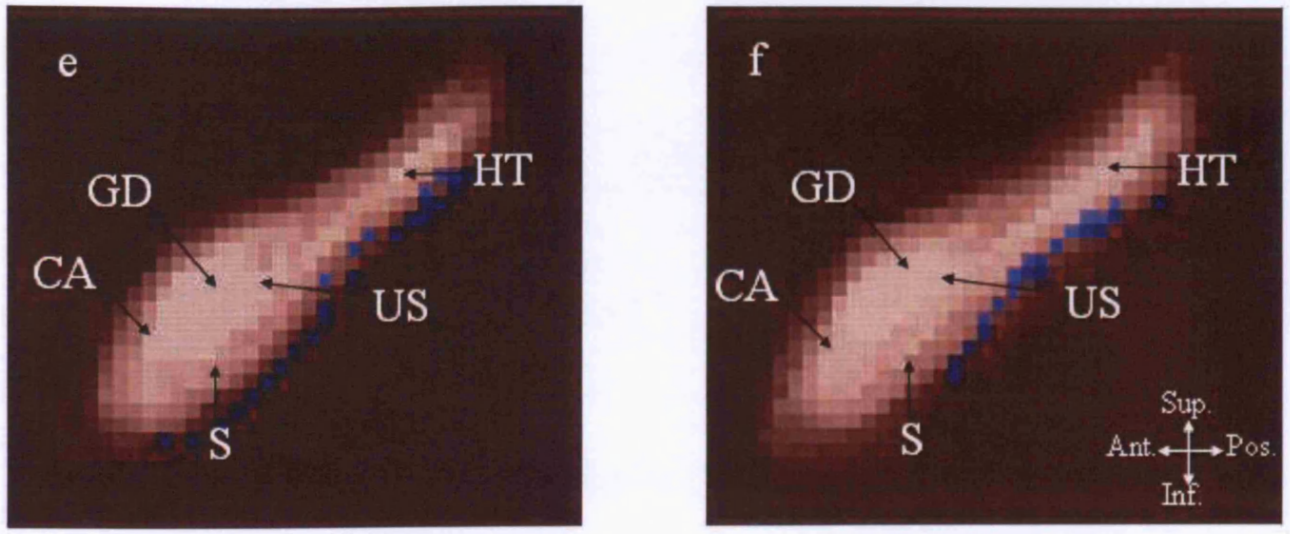


Figure 10.11 Sagittal images of control average maps overlaid with Fisher's exact test results showing where AD significantly > controls at baseline imaging.
Images taken of both the right (e) and left (f) hippocampus. The overlay shows $p \leq 0.05$ where blue and $p > 0.05$ where red.
Anatomical labels: CA: cornu Ammonis, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus, HT: hippocampal tail.

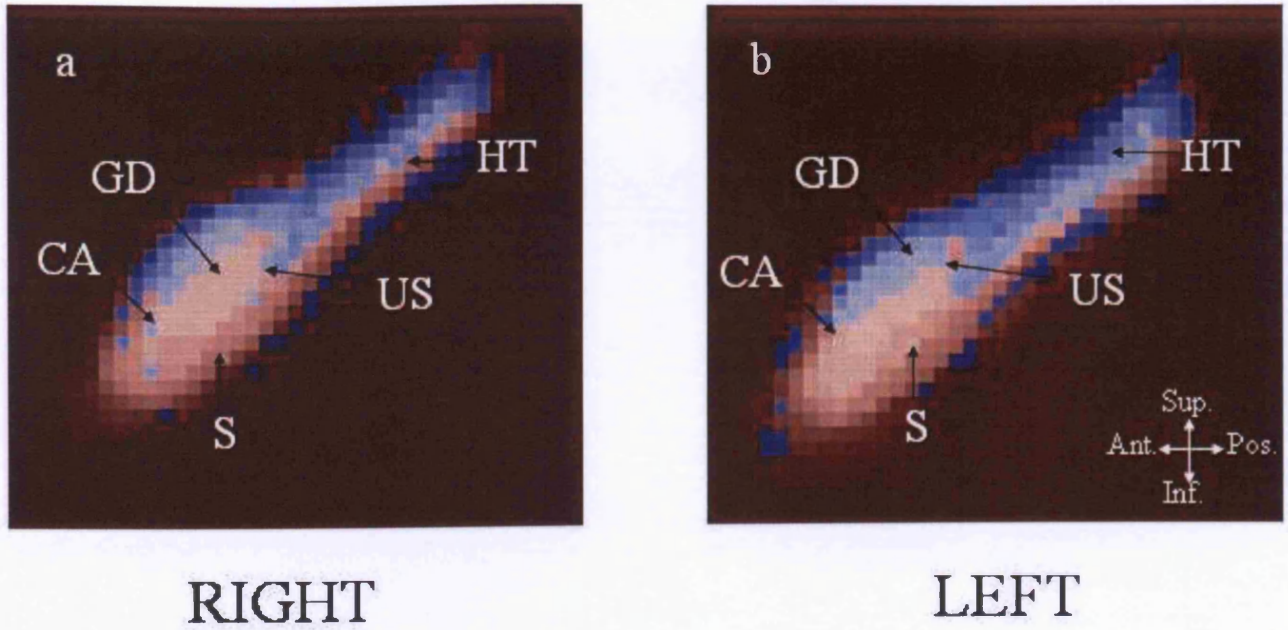


Figure 10.12 Sagittal images of control average maps overlaid with Fisher's exact test results showing where AD and controls are significantly different at repeat imaging.
Images taken of both the right (a) and left (b) hippocampus. The overlay shows $p \leq 0.05$ where blue and $p > 0.05$ where red.
Anatomical labels: CA: cornu Ammonis, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus, HT: hippocampal tail.

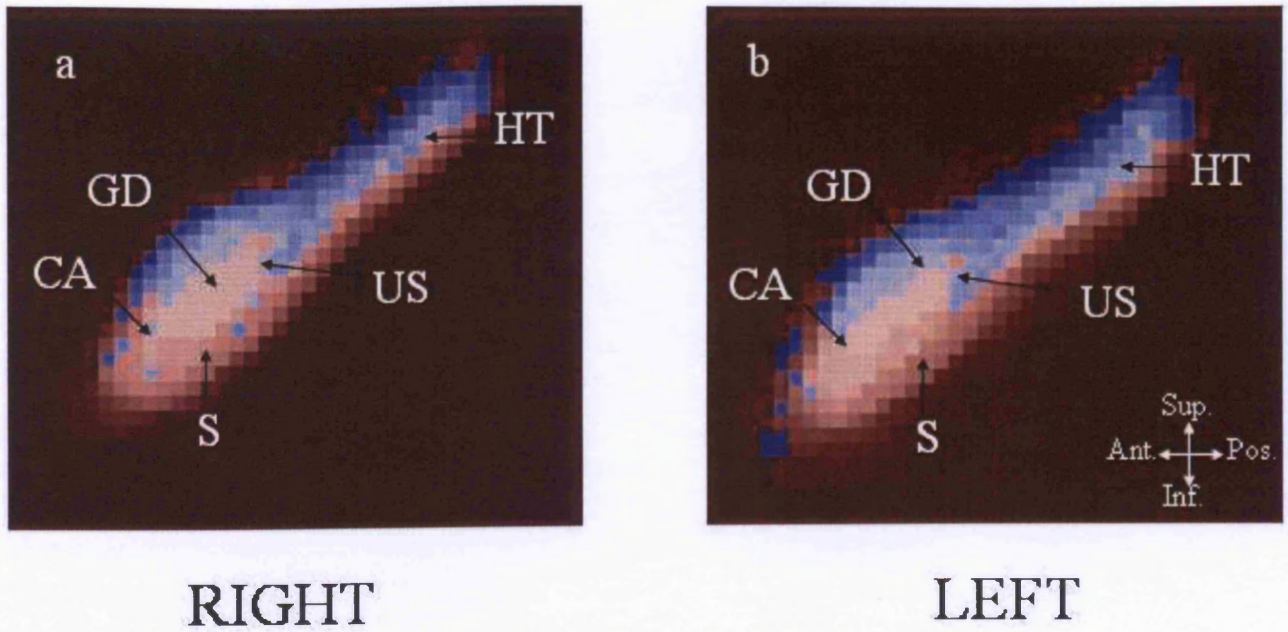


Figure 10.13 Sagittal images of control average maps overlaid with Fisher's exact test results showing where AD significantly < controls at repeat imaging.
Images taken of both the right (a) and left (b) hippocampus. The overlay shows $p \leq 0.05$ where blue and $p > 0.05$ where red.
Anatomical labels: CA: cornu Ammonis, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus, HT: hippocampal tail.

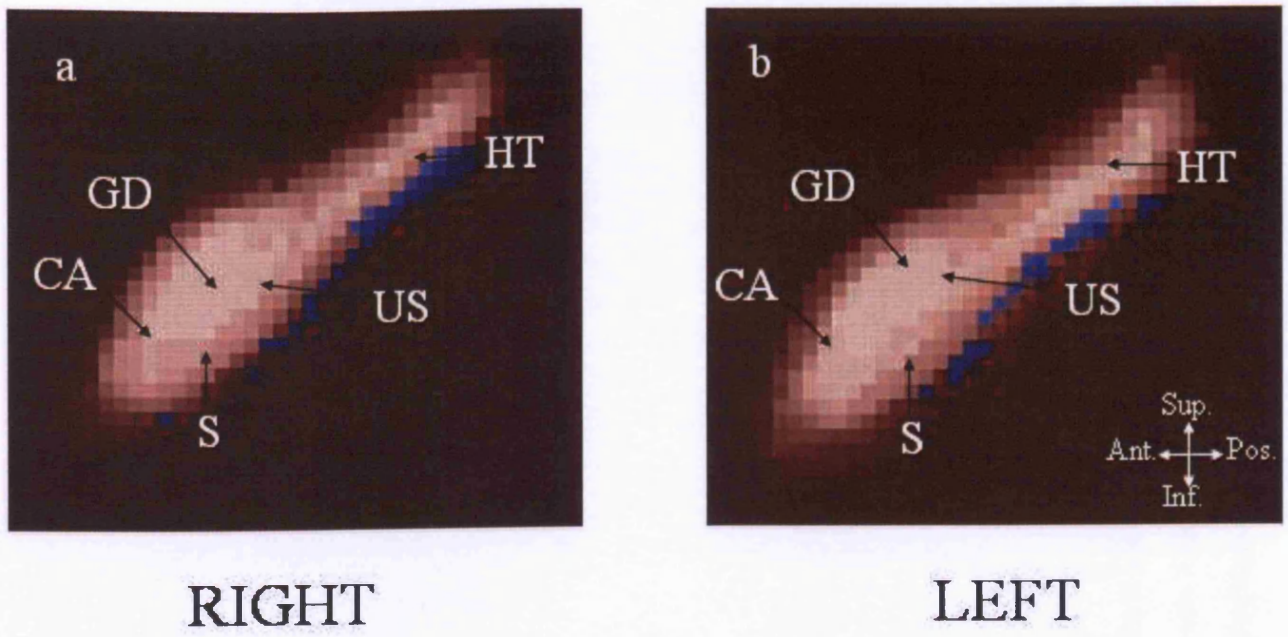


Figure 10.14 Sagittal images of control average maps overlaid with Fisher's exact test results showing where AD significantly > controls at repeat imaging.
Images taken of both the right (a) and left (b) hippocampus. The overlay shows $p \leq 0.05$ where blue and $p > 0.05$ where red.
Anatomical labels: CA: cornu Ammonis, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus, HT: hippocampal tail.

10.3.2. Longitudinal analysis

Figure 10.15 and Figure 10.16 show the difference images between time-points of the control group in both left and right hippocampi. Figure 10.17 and Figure 10.18 show the difference images between time-points in the AD group in both left and right hippocampi.

When applying McNemar's tests to these longitudinal data, no voxels were shown to be significantly different in terms of whether they were labelled as hippocampi on either the left or right and in either the AD or control groups. As a result no images of these statistical maps are presented.

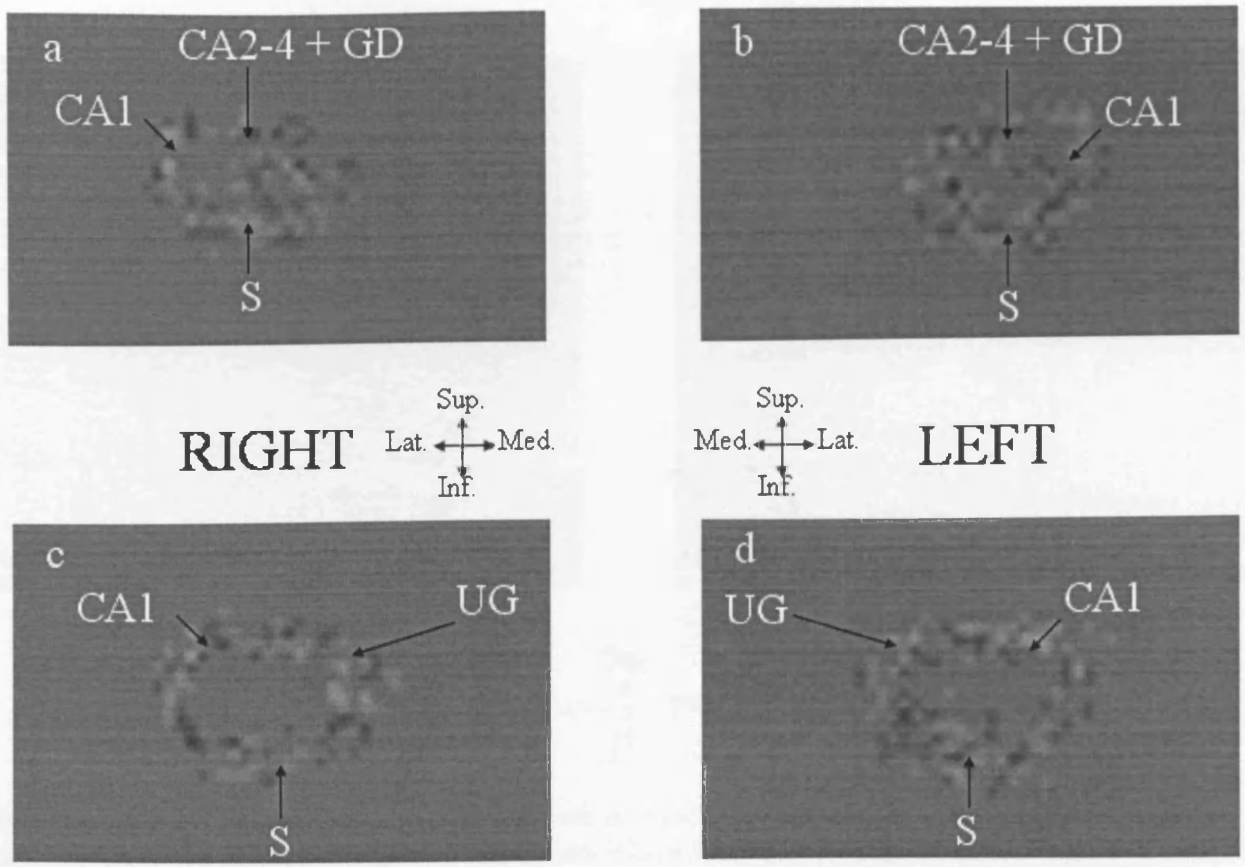


Figure 10.15 Coronal difference images of hippocampal maps of controls at baseline vs. repeat imaging.
Images taken through hippocampal body (a and b) and hippocampal head (c and d) in the right (a and c) and left (b and d) hippocampus. White shows where baseline < repeat and black where baseline > repeat, grey shows no difference between the two time-points.
Anatomical labels: CA1: cornu Ammonis section 1, CA2-4: cornu Ammonis sections 2-4, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus.

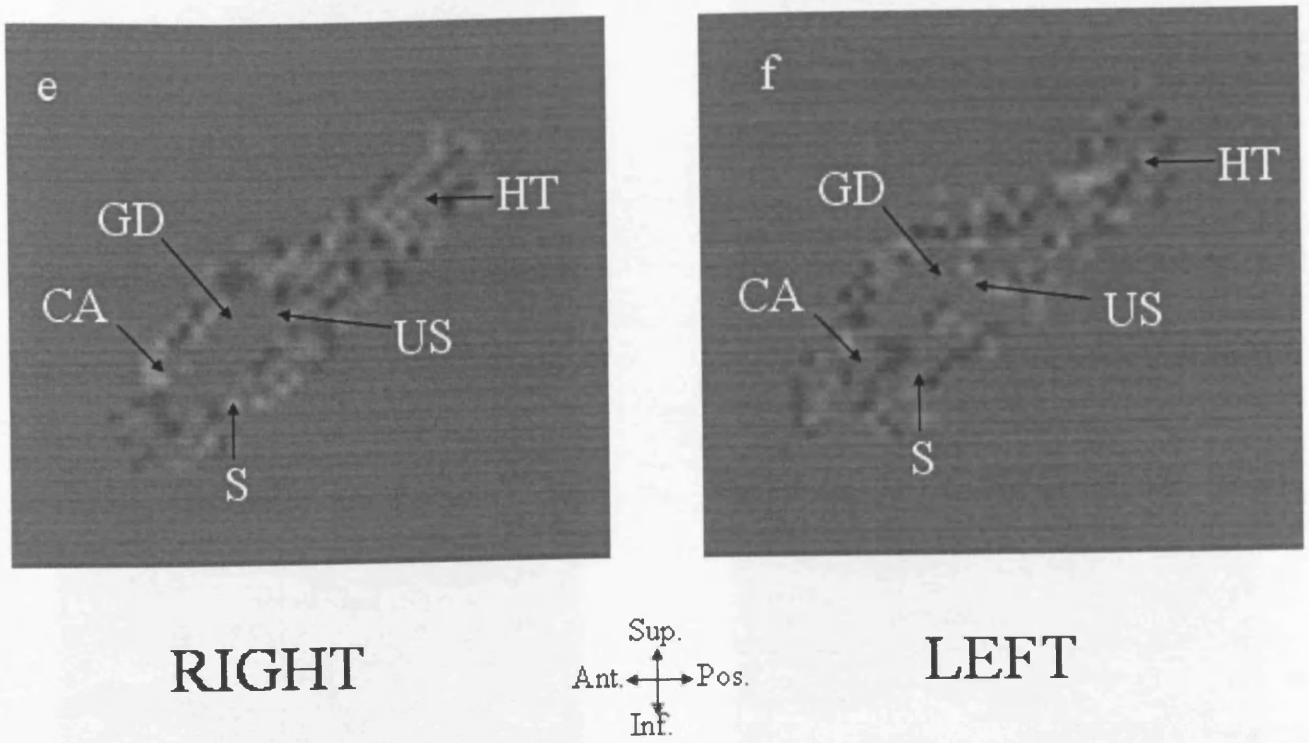


Figure 10.16 Sagittal difference images of hippocampal maps of controls at baseline vs. repeat imaging.
Images taken through the right (e) and left (f) hippocampus. White shows where baseline < repeat and black where baseline > repeat, grey shows no difference between the two time-points.
Anatomical labels: CA: cornu Ammonis, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus, HT: hippocampal tail.

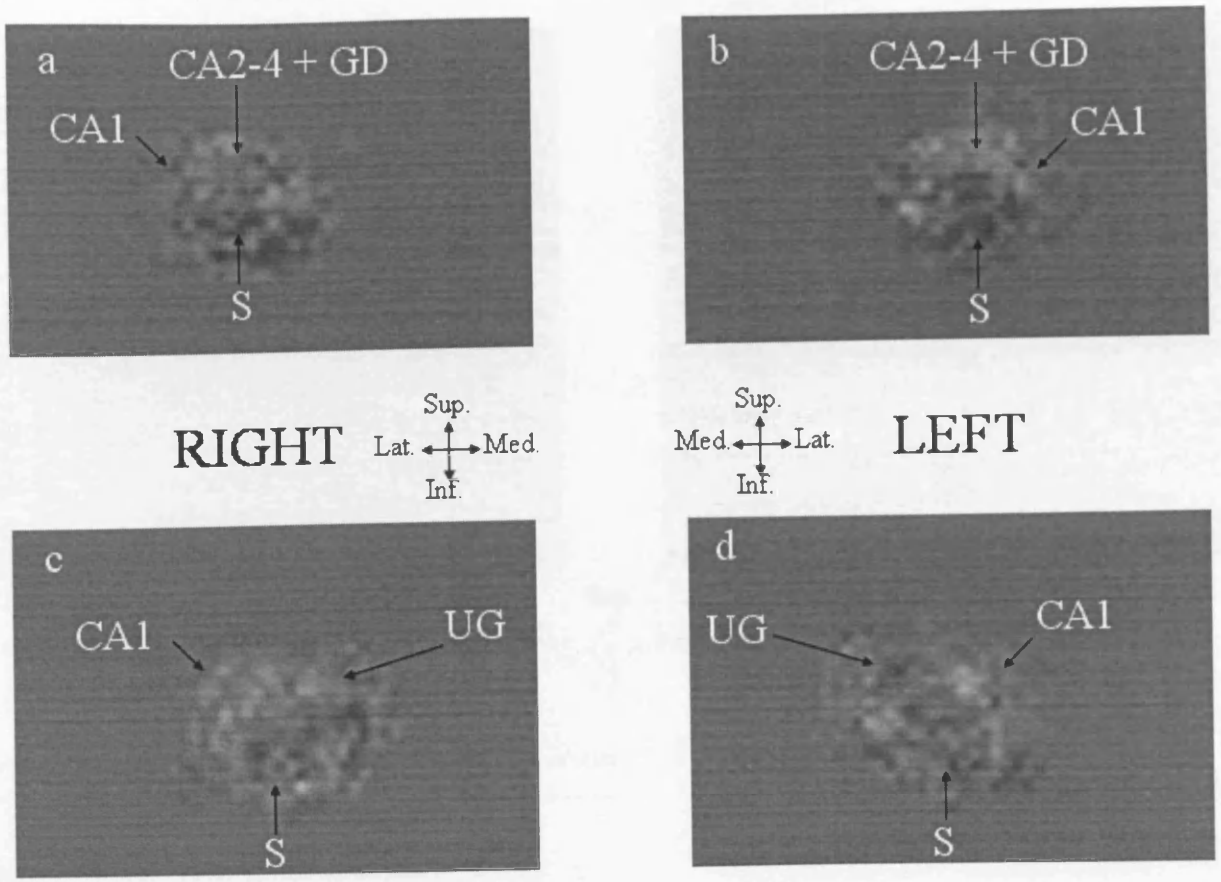


Figure 10.17 Coronal difference images of hippocampal maps of AD subjects at baseline vs. repeat imaging.
Images taken through hippocampal body (a and b) and hippocampal head (c and d) in the right (a and c) and left (b and d) hippocampus. White shows where baseline < repeat and black where baseline > repeat, grey shows no difference between the two time-points.
Anatomical labels: CA1: cornu Ammonis section 1, CA2-4: cornu Ammonis sections 2-4, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus.

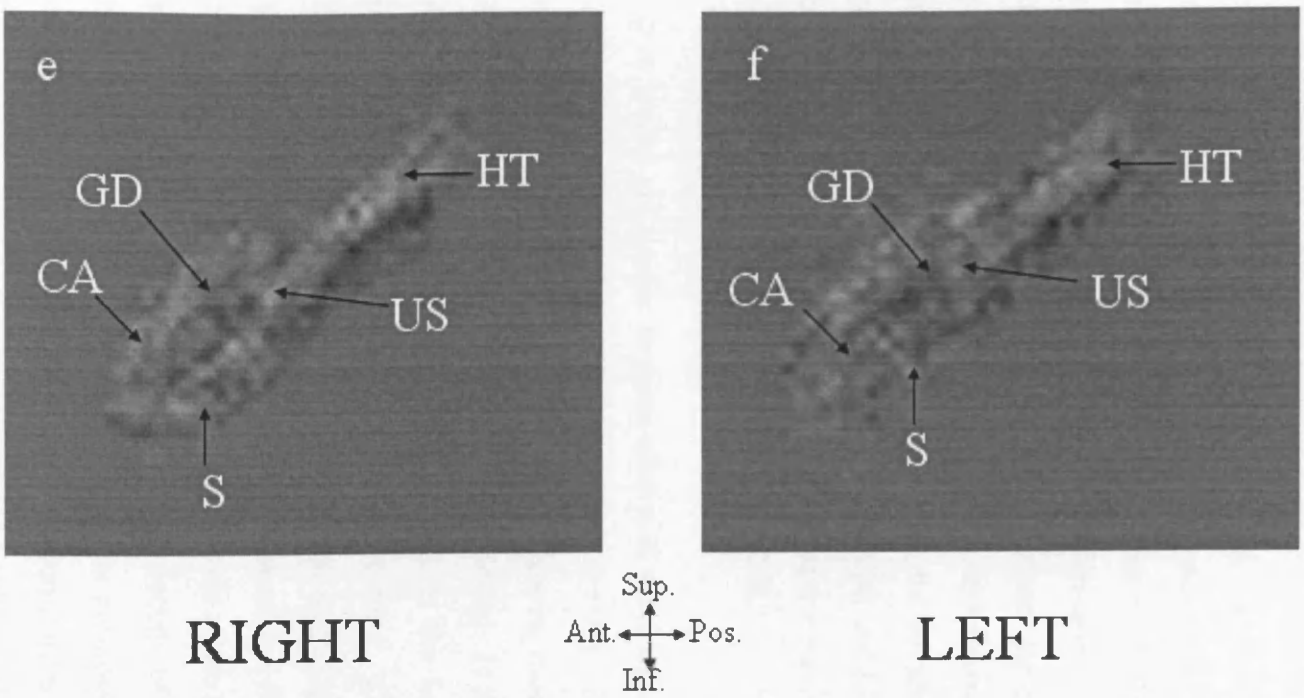


Figure 10.18 Sagittal difference images of hippocampal maps of AD subjects at baseline vs. repeat imaging.
Images taken through the right (e) and left (f) hippocampus. White shows where baseline < repeat and black where baseline > repeat, grey shows no difference between the two time-points.
Anatomical labels: CA: cornu Ammonis, GD: gyrus dentatus or dentate gyrus, S: subiculum, UG: uncinate gyrus, US: uncal sulcus, HT: hippocampal tail.

10.4. Discussion

Cross-sectional findings

The cross-sectional analyses in this study showed differences between AD and control hippocampi at both baseline and repeat time-points. The difference images of the average maps (Figure 10.2 - Figure 10.5) allowed an assessment of the localisation of these differences. Specific areas that were consistently implicated as showing AD < controls include CA1-4, the gyrus dentatus, the uncinate gyrus and the uncal sulcus. The dark areas in the inferior of the subiculum represent areas where controls < AD subjects, suggesting a downward displacement of the hippocampus in AD subjects which was not removed by linear transformations. When testing for significant differences between subject groups at baseline and repeat imaging, the areas which showed significant differences (AD < controls) approximated to CA1-4, gyrus dentatus, and the uncinate gyrus. The uncal sulcus did not show to be significantly different, perhaps owing to the variability in the position and labelling of this feature. As could be inferred from the difference images, the area where significant differences were seen (controls < AD) approximated to the subiculum.

These results are similar to those shown in previous studies (Csernansky *et al.*, 2000; Csernansky *et al.*, 2005; Thompson *et al.*, 2004a). One study showed differences at both baseline and follow-up (1.5 years following baseline scanning) in the groups of AD and control patients (Thompson *et al.*, 2004a). In this study the differences between AD and controls were principally located in the head of the hippocampus which includes the dentate gyrus and subregions CA1-4. Csernansky's group found that in their subjects (18 mild AD and 18 controls, mean ages 74 years) the location of the atrophy was in the head and lateral aspects of the hippocampus. In this study they differentiate between expansion and contraction of the fitted surfaces (similar to the left- and right-tailed tests reported in the results section of this chapter) (Csernansky *et al.*, 2000). However, the areas where the results presented as part of this chapter showed AD to be greater than controls (inferior regions) are not displayed as part of that paper. Csernansky's study concludes that the CA1 subfield is most affected by early AD. The methods employed by these studies differ from those reported in this chapter as they utilise surface-based registration techniques which generate correspondence between points of the mesh framework in all hippocampi in a group of subjects. At the stage of generation of surfaces, the binary masks created in the first instance were smoothed which allows the borders of hippocampi to go through voxels. As a result of this, these

methods may be more sensitive to smaller deformations of the hippocampus than the technique reported in this study.

Pathology remains the gold-standard in determining which sub-regions of the hippocampus may be affected by a disease process. However, pathological load may not directly relate to volume loss which can be measured using MRI. A number of studies have been conducted assessing the pathological load and location in Alzheimer's disease, in addition to neuronal counts. One study assessed cognitively normal people, determined at date of recruitment, over time until death. Some of these subjects over the period of clinical assessment became demented. Pathological examination of the brains revealed that those who were demented had higher amyloid lesion load and presence of NFTs in the neocortex, with NFT load associated with MMSE. Presence of NFTs and amyloid plaques in those who were not classified as demented had the predilection for limbic areas, suggesting that lesion load in areas such as the hippocampus may not be related to presence of dementia (Green *et al.*, 2000). Other studies which have assessed sub-regions of the hippocampus have shown some varied results: amyloid in the subiculum being a predictor of the stage of dementia, but the amyloid being heterogeneous in both location and volume (Bussiere *et al.*, 2002); NFT and neuron counts in the CA1 subregion being associated with MMSE but amyloid not being a good predictor of MMSE (Giannakopoulos *et al.*, 2003); and that variability in CDR is best explained by NFT numbers in the CA2 region and dentate gyrus, not by neuronal loss in the CA1 region (Von Gunten *et al.*, 2005). Although these studies seem to reveal conflicting results between the association of pathology and stage of disease, one of the latter three studies found that volumes in the CA1 and subiculum were reduced in AD (CDR > 0.5 -5) compared with controls (CDR 0 – 0.5) (Bussiere *et al.*, 2002), and another showed a reduction in neuron numbers in the CA1 region in AD (CDR 0.5-2) compared with controls (CDR 0) (subiculum not measured in this study) (Von Gunten *et al.*, 2005). The results presented in this chapter agree with the measures reported in these papers in terms of areas affected apart from our findings in the subiculum as this structure may not have been optimally aligned. In addition, changes in the uncinate gyrus were found which were not analysed as part of the pathological studies mentioned above.

Some areas of the hippocampus were shown to be greater in AD than controls. These areas were located along the inferior part of the hippocampus in the subiculum. This

result may reflect mal-alignment of the hippocampi or a shape change in the hippocampus of AD patients which allows the subiculum to become more curved. Mal-alignment may be caused by other areas of the hippocampus having more of an impact on the registration, i.e. having a greater influence on the cost function, and as a result the inferior border of the subiculum may become poorly registered. A shape change in the inferior subiculum grey / white matter border may be due to the fact that as the hippocampus and its supporting structures in the medial temporal lobe become atrophied this causes a structural change such that the whole parahippocampal gyrus dips and curves more than in those subjects in which no atrophy has occurred.

Longitudinal findings

The longitudinal analysis showed no areas of loss were detected in particular regions of the hippocampus over time. One other study has found differences over time in a relatively large group of early AD patients (MMSE 25) and age matched controls (Wang *et al.*, 2003). In that study paired Wilcoxon's sign rank tests showed significant areas of inward deformation in the early AD cases and ascribed this to specific losses in the CA1 subfield of the hippocampus (see Figure 8.10 for labelling of hippocampal subfields). That study differed from the study reported here in terms of the methods employed to detect change. Much like Thompson's group, surface-based registration and therefore correspondence and deformations were obtained. In addition to this the scanning interval was over two years (twice as long as the scanning interval used to generate longitudinal results in this chapter) making the changes in the hippocampus greater and potentially easier to detect. Using the techniques reported in the methods section of this chapter, the hippocampus would have to shrink by at least a voxel in places surrounding the hippocampus and these areas would have to be well aligned in order to generate significant longitudinal patterns of atrophy.

General discussion

Differences between the study presented in this chapter and those of others are not only methodological. Differences in the cohorts in terms of stage of disease or disease progression may account for varying results. Age might also contribute to the differences seen. Heterogeneity of the distribution of pathology caused by AD as demonstrated by pathological studies may also influence where differences in the hippocampus can be seen (Bussiere *et al.*, 2002).

This study is limited by a number of factors. First the interpolation scheme chosen to generate hippocampi from their original space to template space may not be optimal. Trilinear interpolation was used for this procedure but it may be that another method provides more accurate transformation of the hippocampal regions. The registration algorithm used to align the hippocampi into template space may also require investigation. Affine brain-brain registration was used to obtain an approximate spatial match followed by rigid-body hippocampus to hippocampus registration. It may be that non-linear registration with increased degrees of freedom would match the hippocampi more accurately; however such registration might also remove the differences between hippocampi the experiment is designed to detect. The technique of registering brain to brain in the first instance may not optimally correct for head-size. This may be important since some of the variance in the hippocampal regions, which will affect the p values obtained, may be accounted for in terms of head-size. Lastly, correction for multiple comparisons was not performed following generation of the statistical maps. Essentially, performing many statistical tests over the whole hippocampal region is likely to lead to detection of false-positive voxels and the results presented here did not correct for this potential source of error.

One potential problem that this type of research highlights is that such studies rely on inter-subject correspondence which is assumed rather than necessarily obtained. Such correspondence is hard to measure and assess and as a result it is difficult to establish that one area of the hippocampus is differentially affected by the disease. This may be the result of the registration procedure finding better alignment of one particular region rather than true correspondence. For example, it may be that the inferior boundary of the hippocampus in the head region gives the best cost function results following registration, leading to the differences being found at the superior border when this region is not necessarily the area from which the tissue has been lost. Surface-based techniques usually rely upon correspondence between parametric meshes created from the binary masks. This type of matching assumes that point A on one hippocampus relates to point A' on another hippocampus, whereas it may be that point A relates to 80% point A', 10% point B' and 10% point C'. Such assumptions may influence results found by these techniques.

10.5. Chapter conclusion

In conclusion, differences between AD and controls in the location of regions labelled hippocampus by manual segmentation methods can be seen cross-sectionally. This study highlights areas which may be preferentially affected by AD which include the uncinate gyrus, CA regions 1-4 and the dentate gyrus. However, this study also highlights issues of aligning structures such as the hippocampus between subjects which can lead to assumptions as to which areas of the hippocampus may be preferentially affected by the disease.

11. APPLICATIONS OF SEMI-AUTOMATED TECHNIQUES TO A CLINICAL TRIAL

11.1. Chapter introduction

Background

Previous chapters have demonstrated the utility of hippocampal rates of atrophy in diagnosis of disease (Chapters 5 and 6, pages 134 and 162), detection of early change (Chapter 6, page 174), and their potential in tracking the progression of disease (Chapter 6, page 174 and Chapter 5, page 153). Chapters 8 and 9 demonstrated automation of these measures in a controlled situation where the acquisition and subjects were selected by the research group performing the atrophy measures. Two studies featured earlier have been based at one site only (see Chapter 5, page 153, and Chapter 9, page 243), the remainder of the studies utilised scans from a small number of sites (<4). This type of analysis is useful for establishing the concept of methods for determining the rates of atrophy, however there is a risk that the methods developed from such studies are not robust to problems associated with scanning at multiple sites.

Few clinical trials or large studies have been conducted reporting hippocampal rates of atrophy as an outcome measure. One reported on the results of a trial of a melameline (a muscarinic receptor agonist). This therapy was proposed to augment diminished cholinergic function which occurs in AD. The trial was suspended due to lack of efficacy but 100 of the treated and 92 of the placebo subjects completed the MR analysis part of the study. This study showed MR results to be reliable across sites with decline more consistently seen with MR rather than clinical measures (Jack *et al.*, 2003).

Another trial involved active immunisation of subjects against one possible pathological substrate of AD: aggregated Aß protein (see Chapter 1, page 20). This trial showed manually-derived hippocampal rates of atrophy to be greater in those treated compared with those in the placebo arm, and although this difference was not significant in the hippocampus, there were significant differences in both rates of ventricular expansion and brain shrinkage (higher in the treatment arm) (Fox *et al.*, 2005). The trial was suspended due to adverse events in the treatment arm, with encephalitis being detected in a number of cases (Orgogozo *et al.*, 2003).

One other trial investigated the effect of an acetylcholinesterase inhibitor (donepezil) on MR-based markers in AD (Krishnan *et al.*, 2003). This pilot study ran for 24 weeks and found that the treated patients had reduced loss of the hippocampus compared with the placebo arm. However, the results for this study were surprising in that the rate of change seen over the 24 weeks in the placebo arm would approximately equate to over a 16% hippocampal loss per year. In addition, the results in the treatment arm were asymmetric; the left and right hippocampi had changes of equal magnitude, but in different directions (i.e. loss was seen in the left and gain on the right).

In order to establish whether the techniques for assessment of hippocampal atrophy rates described earlier have utility in clinical trials, it is important to test them in a multi-centre setting.

Chapter outline

This chapter aims to evaluate the variability of hippocampal rates of atrophy generated by a semi-automated measure (HBSI) in and the discrimination between the placebo and treatment groups in a clinical trial setting. The data used is from the AN1792 trial of Aß immunisation (Fox *et al.*, 2005) where manual measurements were already available as a comparison.

11.2. Methods

11.2.1. Patients

Patients and sites are described in detail elsewhere (Fox *et al.*, 2005; Gilman *et al.*, 2005). Patients included in the trial were aged 50 to 85 years, met the criteria for a diagnosis of probable AD as defined by the NINCDS-ADRDA (see Appendix One) and had an MRI brain scan supporting the clinical diagnosis of AD. This study assessed the differences in those who were treated with the drug and responded to treatment (this was defined as those subjects who had an antibody response of a serum anti-AN1792 IgG (total) titre of ≥1:2,200 at any time after injection 1 and included those subjects that developed encephalitis) and those subjects in the placebo group (see Table 11.1).

Table 11.1 Subject demographics.

	Placebo	Responders
Number	57	45
M/F	23/34	24/21
Mean age (SD), in years	70.7 (8.2)	72.2 (7.1)
Mean (SD) MMSE	20.2 (3.5)	20.8 (3.5)
Mean duration of AD years (SD)	3.9 (1.8)	3.8 (2.1)

11.2.2. Trial details

The study was conducted at 28 centres in five countries between September 2001 and December 2002. Patients were randomly assigned in a double-blind manner to receive treatment with AN1792 225 μg (Elan Pharmaceuticals, South San Francisco, CA) and QS-21 50 μg (Antigenics, Framingham, MA) containing 0.4% polysorbate 80 or normal saline in a 4:1 ratio. The assignment code was held by an independent statistician. Patients received only one to three (day 0, months 1 and 3) of the planned six injections because of the premature discontinuation of the study because of the development of encephalitis in 6% of the treated patients. All patients who had been enrolled in the study could participate in the safety follow-up period, occurring for at least nine months after their last dose of study treatment.

11.2.3. Image acquisition

Images were collected at 17 scanning sites owing to pooling of the 28 clinical assessment centres. The study outcomes included measures derived from volumetric MRI brain scans performed at baseline and at month 12 or early termination. Patients were imaged on 1.5 T MR units using a standardised protocol; all efforts were made to ensure each follow-up scan was performed on the same scanner using the same imaging protocol as the baseline assessment.

Volume measurements were derived from a T1-weighted three-dimensional gradient echo sequence. Exact acquisition parameters varied depending on the scanner manufacturer (see supplementary information from (Fox *et al.*, 2005)); all were coronal volumetric acquisitions lasting 7.5 minutes (to minimise movement artefact), with the slice (partition) thickness of 1.5 to 1.8 mm adjusted to cover the entire brain, a within-plane field of view of 25 x 25 cm, and an effective matrix size of 256 x 256 x 124. At the baseline visit, each site also performed standard imaging (including T2-weighted or fluid-attenuated inversion recovery sequences) to exclude non-AD pathology.

11.2.4. Image processing

Prior to hippocampal segmentation, all scans had brain regions segmented and subsequently were rigidly registered to the MNI 305 brain template to ensure similar orientation of scans. Scan pairs were co-registered using nine degrees of freedom (dof). Hippocampal volumes were manually delineated from each individual's scan pair. Left and right hippocampi were delineated as described in Appendix Four using "mirror-image" volumetry (see Chapter 4 page 101). One protocol deviation necessary for this trial was the application of 60% as the lower threshold rather than the usual 70% mean brain intensity. This was due to differences in scanning acquisition which meant the 70% caused some hippocampal tissue to be excluded from the hippocampal region. Segmentations were performed by three operators. The operators were blind to all patient information and other measurements at the time of segmentations. Tracing of the hippocampal boundaries was performed with each scan of the registered pair viewed side by side to improve consistency. Each subject (baseline and repeat image, left and right hippocampus) was segmented by only one segmentor. All measurements were performed blinded to treatment group, time-point or subject details. The detailed results of the MRI measures are reported and discussed in (Fox *et al.*, 2005). In summary, the subjects who developed a predetermined titre of antibodies (responders) had increased brain volume loss (3% vs. 2%) compared with the placebo group. Manual hippocampal volumetry showed a non-significant loss (3 (3)% vs. 4 (3) % loss in placebo and responder groups respectively).

11.2.5. HBSI calculation

Following completion of the trial, information regarding each subject's scan dates, treatment group assignment and antibody titre were given to the Dementia Research Centre. HBSI was then calculated in a similar manner to that described in Chapter 8, page 205. In brief, the manually-drawn baseline hippocampi were used and the HBSI was calculated on DBC-corrected nine dof registered scan pairs. The window widths and window centres that were selected for this dataset were 0.2 and 0.55 respectively.

11.2.6. Statistical analysis

Mean rates of atrophy were calculated for total (left plus right) hippocampi. Atrophy rates were analysed on a logarithmic scale [(log (follow-up volume/baseline volume))/interval] in order that doublings and halvings in volume be treated as effects of equal magnitude. Follow-up volume with the HBSI technique was calculated as baseline volume – HBSI measure. Mean atrophy rates were calculated by back transformation with SDs calculated from variance transformation formulae. Tests of means and variances were performed on a log scale. Paired t tests were used to assess whether the mean rates of atrophy differed according to method within each subject group. Pitman's tests were used to assess whether variances differed between methods within each subject group. Student's t tests (allowing for unequal variances) were used to assess whether mean rates of atrophy differed between each group according to method. Linear regression was used to assess the association between the two methods in each subject group.

11.3. Results

The mean (SD) rates of atrophy (% per year) for all responder and placebo subjects were 3.61 (3.25) for manual and 3.42 (3.60) for HBSI. Mean (SD) rates of atrophy by group are displayed in Table 11.2 and Figure 11.1. As shown in Table 11.2, there was evidence of a difference between responders and placebo using the HBSI which was not shown using manually-derived rates of atrophy. HBSI and manual measures were not significantly different in the placebo group in terms of means (p=0.185) or variances (p=0.163). There was some evidence that there was a difference between the methods in both means (HBSI > manual, p=0.022) and variances (HBSI > manual p=0.007) in the responder group. Removing encephalitis cases from the responder group, manual and HBSI generated rates of atrophy gave mean (SD) rates of atrophy of 3.94 (3.02) and 4.73 (3.69) respectively, (p=0.077).

Figure 11.2 shows the association of rates of atrophy according to each method within each subject group. Those responders who developed encephalitis are shown as triangles. There was some association between the two methods with R^2 being 0.47 ($p<0.0001$) in the responders and 0.25 ($p=0.0001$) in the placebo groups.

Table 11.2 Mean (SD) atrophy rates in placebo and responder groups (%/year).

	Manual	**HBSI**
Placebo	3.23 (3.54)	2.64 (3.02)
Responders	4.08 (2.82)	5.08 (3.82)
p value	p= 0.177	p < 0.001

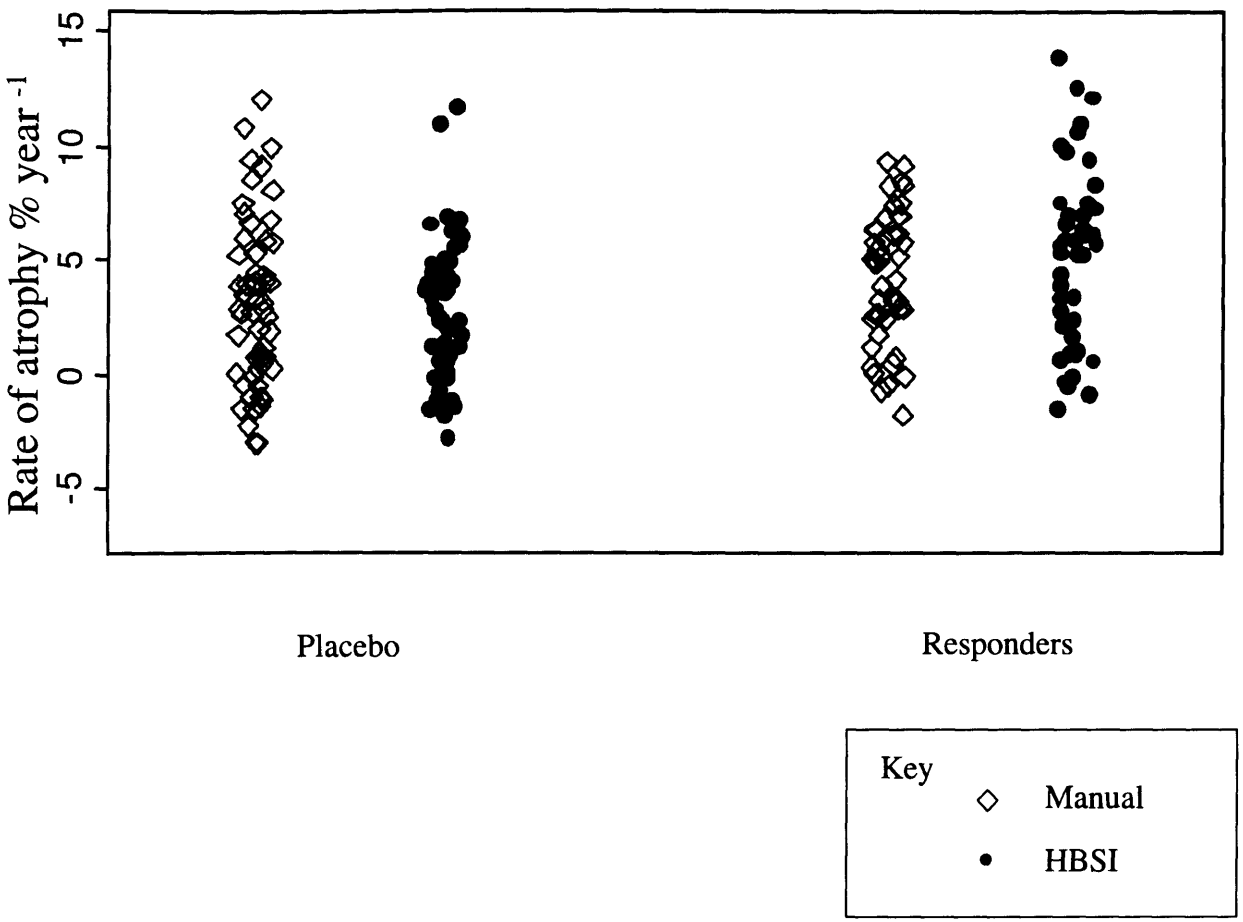


Figure 11.1 Annualised rates of atrophy of the placebo and responder group using both manual and HBSI methods.

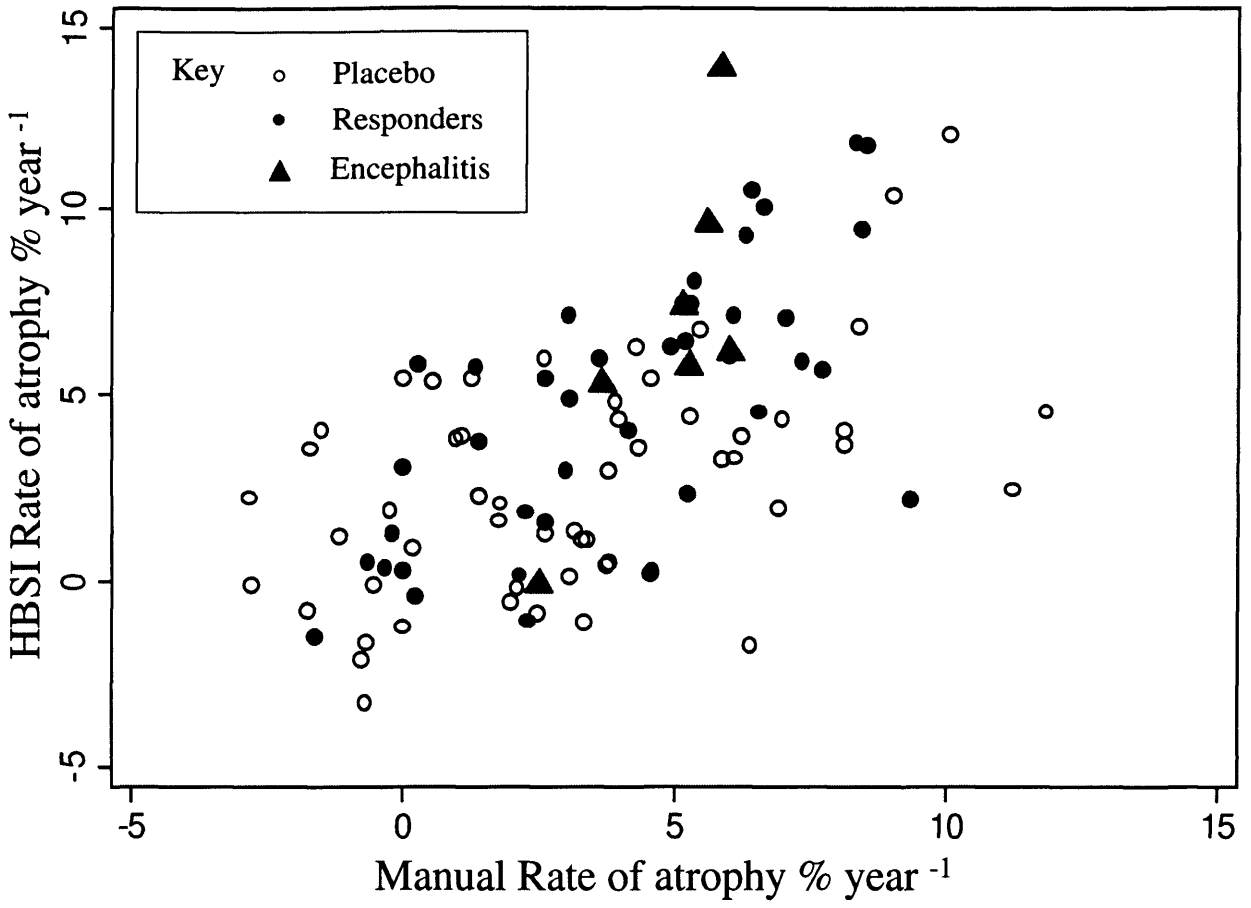


Figure 11.2 Association of HBSI and manual rates of atrophy (both annualised) according to subject group. Encephalitis cases form part of the responder group but these are highlighted in red.

11.4. Discussion

This study has shown that HBSI gives consistent results in a clinical trial situation (17 sites). The fact that a significant difference could be detected between subject groups using HBSI but not manual measures is due to the mean rate being lower in the placebo group (although not significantly) and corresponding rate in the treatment group being significantly higher. A similar comparison of means between subject groups using manual measures did not reveal a significant difference. This may be due to a number of factors. First the HBSI was calculated using differentially bias-corrected images owing to the inconsistent bias field between baseline and repeat scans, whereas manual measures were performed on image pairs which were not corrected in this way. Measuring hippocampi on these images following DBC may have also revealed significantly different results between subject groups. Secondly, HBSI may reduce some of the variance caused by operator error as discussed in Chapter 8, page 219. Removal of the encephalitis cases reduced the difference between mean atrophy rates generated from HBSI and manual measures, inferring that the greater atrophy rates calculated in the encephalitis cases using HBSI may have contributed to the overall higher rate in the responder group.

Although the HBSI detected a significant difference between the two subject groups, there remained a great deal of overlap in the results of the two subject groups. This is perhaps unsurprising since rates of atrophy in AD subjects have been shown to be heterogeneous in studies described in this thesis (see Chapters 5,6,8,9, pages 143, 156, 166, 174, 178, 212, 230, and 252). The rates of atrophy calculated as part of this study are similar to other studies described and Table 1.3, but in the placebo group, are slightly lower than other mean rates calculated in this thesis. The rates of atrophy in the placebo arm are also lower than the placebo arms of both of the other two multi-centre trials which have used hippocampal atrophy rates as an outcome measure. In the melameline trial, median rates of atrophy were approximately 5% (Jack *et al.*, 2003). In the donepezil trial, the inter-scan interval was relatively short (24 weeks) and the change calculated over that period was 8.2% (Krishnan *et al.*, 2003). However these changes roughly equate to more than a 16% annualised rate of hippocampal atrophy which would be higher than any other annualised rate reported (see Table 1.3).

The lower rates of atrophy quoted in this trial may be due to the range of sites and scanners used in this trial. This may also be caused by the application of DBC during the brain-brain registration procedure. DBC has been shown to cause slight underestimation of the subsequent atrophy measure (BSI) since some of the atrophy is modelled as part of the differential bias. However, this underestimation may occur in all subjects irrespective of group, and the benefits of using such correction may outweigh the cost. As a result the relative difference between groups may be a more important assessment. Further work is required to assess the effect of DBC on the BSI of small regions such as the hippocampus.

Although significant differences between the two groups were shown using the HBSI it is important to note that change in the other regions (brain and ventricular BSI) also showed significant differences. HBSI does not reveal a group difference which has not been shown in these subject previously using different structures (Fox *et al.*, 2005).

The treatment administered as part of this trial was intended to raise an immune response against one of the supposed pathological substrates of AD (Aß amyloid protein). The exact cause of the increased rates of atrophy seen in the responder group remains unexplained. One theory which may partly explain this increased rate of atrophy is that the protein is removed from the brain and as a result the brain decreases in size by that volume. Some supporting evidence for this theory is derived from reports of patients who were immunised and had subsequently died. Post-mortem of their brain material suggests widespread and sometimes patchy removal of the Aß amyloid protein (Ferrer *et al.*, 2004; Masliah *et al.*, 2005; Nicoll *et al.*, 2003). Post-mortem measures of the volume of amyloid in the hippocampus have been calculated to be approximately 1% in AD cases (calculated from sampling of tissue in CA1-3 subregions and the dentate gyrus in 90-100 year old subjects with a CDR of 1-2) (Von Gunten *et al.*, 2005). A similar study in younger subjects which were more severely affected (CDR 2-5, average age 89 years) showed the amyloid load of the CA1-3, dentate gyrus and subiculum to be approximately 6-7% (Bussiere *et al.*, 2002). Amyloid removal (Nicoll *et al.*, 2003) may therefore partly account for the increased hippocampal loss in the responder group.

11.5. Conclusion

HBSI appears to be a viable measure in a multi-site trial, providing comparable results to those generated by manual volumetry. One advantage of using this technique is that only half of the number of hippocampal segmentations may be required, potentially reducing operator error in addition to reducing the workload substantially.

12. META-ANALYSIS

12.1. Chapter introduction

Background

A relatively large number of studies assessing longitudinal hippocampal change have been reported in the literature (see Table 1.3). Many different methods have been used to generate these rates of atrophy and different populations of patients have also been included in the different studies. To the best of my knowledge, no statistical review of the literature has been conducted to date. Such a review is required to assess heterogeneity of reported studies to allow understanding of the effects of age and disease severity on the calculated atrophy rates, and to pool the results from these studies to estimate the rate of atrophy of the hippocampus in AD and matched control groups. It may also be useful in identifying outlier results where the methodology employed may deserve critical review.

Chapter objective

This chapter focuses on the meta-analysis of hippocampal atrophy rates in patients with AD and matched controls from studies reported in the peer-reviewed literature.

12.2. Methods

The following protocol was produced for conducting the study.

12.2.1. Protocol

General study inclusion criteria

All observational and randomised trials were included, regardless of quality of the study.

Studies in which patients were given symptomatic treatment were included. If the study was a randomised controlled trial (RCT) in which some patients underwent an intervention which was shown to alter the atrophy rate of the whole brain as part of the trial (for example interventions developed to remove the pathological substrate thought to cause the disease or those which may prevent cell death) this arm of the study was excluded, however the placebo arm was included.

Studies in which MRI scans were used were included, but no restrictions were made as to the scanning protocol or strength of magnet used to image the patients or segmentation protocol used to determine the atrophy rate.

Studies reporting results from patients with AD of any age or gender were included. There were no restrictions on the method used to diagnose AD. Unpublished studies were not included and studies where the mean inter-scan interval was less than nine months were also excluded owing to the lack of precision in determining the mean atrophy rates (Schott *et al.*, 2005). As MCI was not a subject group being formally assessed as part of this meta-analysis, cognitively impaired subject groups with an MMSE >26 were also excluded.

Outcome measure

Mean (arithmetic) atrophy rate specified as % loss of baseline volume per year.

Search methods for identification of studies

Searches for relevant studies were performed using electronic and other sources.

1) Electronic sources

PubMed and Google Scholar were searched for studies from any period using the following keywords: hippocampus, AD, atrophy, Alzheimer's, MRI, rates.

2) Other sources

Journals searched included: Neurology, Annals of Neurology, Archives of Neurology, Neurobiology of Aging, American Journal of Neuroradiology, Brain, Cerebral Cortex, Cortex, European Journal of Neurology, Hippocampus, Journal of Neurology, Journal of Neurology, Neurosurgery and Psychiatry, Lancet Neurology, Neurodegeneration, NeuroImage, Neuropathology, Neuroradiology, Practical Neurology, Dementia and Gereatric Cognitive Disorders, Journal of Alzheimer's disease, American Journal of Alzheimer's disease and other dementias, Alzheimer's disease and associated disorders, Lancet, PNAS, Nature.

3) Personal communications

The list of published studies generated were shown to another researcher in this field from a different research group (Laura van de Pol from VU, Amsterdam) who checked whether there were any studies that may have not been included. Once the potential list was finalised between the two research groups, all authors of the studies were contacted to ensure that correct information was being used and any information which was lacking (such as MMSEs, or mean rates of atrophy) was requested. As a result of this request, errors within one published paper were found and an erratum is now in submission (Thompson *et al.*, 2004a). A personal communication with another author highlighted problems with methodology and subsequent results within another paper (Hampel *et al.*, 2005). In addition, the total list was sent to each author to ensure there were no other studies which had been missed by the searches performed. Published studies resulting from searches which are not in English were to be included when possible through contact with authors where necessary.

Quality assessment

Once studies had been selected for the meta-analysis, each was assessed for quality by tabulating the following variables which may explain heterogeneity of the results published in the studies: cohort used (population, case/control, case series), AD diagnostic criteria and exclusion criteria, APOE genotyping, other medications, drop-out rate if RCT, assessment of and exclusion criteria for controls, magnet strength, number of scanning sites, scan "slice" thickness, hippocampal measurement method, number of raters, registered images or TIV corrected, whole hippocampus measured, reliability of method, and blinding of raters to patient information and chronology of scan ordering.

12.3. Statistical analysis

From each study the mean atrophy rate and its standard deviation for the AD patients (and separately for controls where available) was obtained, together with means and standard deviations of the ages, MMSEs and inter-scan intervals in each subject group. Calculation of the standard error required for meta-analysis was made using the standard deviation and the number of patients for all studies.

Data synthesis (meta-analysis)

A random effects meta-analysis was performed, consistent with the belief that each study is estimating a different mean atrophy rate. Restricted maximum likelihood was used to estimate the between-study variance. For studies which reported results separately for different subgroups of AD patients, the means and standard errors were combined to form an overall mean and standard error for the study. Each subgroup was weighted in proportion to the inverse of its mean's standard error, as in a fixed effects meta-analysis. For studies which contained both a control and an AD group, the mean difference in atrophy rates between the two groups was calculated. These differences were then entered in a further random-effects meta-analysis in order to estimate the overall mean within-study difference in rates between AD and control groups.

Assessment of heterogeneity

The extent of heterogeneity between the studies was assessed through examination of the between study variance which was estimated in the random effects meta-analysis. The I^2 statistic was calculated to estimate the proportion of heterogeneity which was not attributable to chance.

Assessment of reporting bias

Due to the focus of this meta-analysis, it was assumed that studies would not be published on the basis of the estimated size of mean atrophy in AD patients. However, this was examined with funnel plots.

Subgroup analysis and investigation of heterogeneity

Heterogeneity in atrophy rates between studies may be explained by the mean age or disease severity (MMSE score) of study participants, or inter-scan interval, through use of meta-regression. It was recognised that estimation of such effects may be limited if only a small number of studies are available.

The ability to estimate the effect of such an explanatory variable on heterogeneity relies on there being sufficient variability in the variable between studies. For the studies which reported results for separate subgroups of AD patients, these subgroups were entered in the meta-regression separately. However, since the groups originate from the same study, they will have many characteristics in common. This was accounted for through the sharing of the same study-specific random effect.

12.4. Results

The resulting studies included in this meta-analysis are displayed in Table 12.1. The quality of these studies is reported in Table 12.2, and those studies which were excluded are detailed in Table 12.3.

Table 12.1 Studies included in the meta analysis.

Study **Author, Year**	**Checked with author**	**Research Group**	**Subject groups**	**n**	**MMSE /30 at baseline** **Mean (SD)**	**Age, years at first MRI** **Mean (SD)**	**Annualised loss of total (left plus right hippocampi)** **Mean % year $^{-1}$ (SD)**	**Interval (years unless specified)** **Mean (SD)**
(Jack *et al.*, 2004)	Yes Means and SDs provided by personal communication	Jack	C stab	40	*29.0 (0.9)* *<34>*	*78.2 (7.7)*	*1.4 (0.7)*	*4.3 (0.7)*
			AD slow	31	*22.3 (4.4)* *<30>*	*73.9 (6.9)*	*3.9 (2.6)*	*1.4 (0.4)*
			AD fast	33	*19.5 (4.0)* *<32>*	*77.8 (8.4)*	*4.8 (3.2)*	*1.4 (0.4)*
(Du *et al.*, 2004)	Yes	Weiner	C	25	29.0 (1.0)	76.8 (7.8)	0.8 (1.7)	2.0 (0.7)
			AD	20	21.0 (7.2)	75.3 (7.2)	5.9 (2.4)	1.8 (0.7)

Study **Author, Year**	**Checked with author**	**Research Group**	**Subject groups**	**n**	**MMSE /30 at baseline** **Mean (SD)**	**Age, years at first MRI** **Mean (SD)**	**Annualised loss of total (left plus right hippocampi)** **Mean % year $^{-1}$ (SD)**	**Interval (years unless specified)** **Mean (SD)**
(Hashimoto *et al.*, 2005)	Yes	Hashimoto	AD no treat	93	21.6 (2.8)	70.5 (9.1)	5.04 (2.54)	392 (35) days
			AD treat	54	21.8 (3.9)	69.5 (9.5)	3.82 (2.84)	388 (29) days
(Barnes *et al.*, 2005)	Yes	Fox	C	50	29.4 (0.8)	59.6 (13.8)	*1.15 (1.86)*	418 (172)
			AD	32	19.4 (5.0)	59.0 (11.3)	*5.49 (3.12)*	456 (311)
(Thompson *et al.*, 2004a)	Erratum in press. Values here are correct.	Thompson	C	*14*	*29.5 (0.9)*	*71.4 (3.2)*	*2.1 (2.8)*	*2.6 (1.0)*
			AD	*17*	*16.8 (6.6)*	*68.7 (6.8)*	*6.8 (10.2)*	*1.3 (0.9)*

Study **Author, Year**	**Checked with author**	**Research Group**	**Subject groups**	**n**	**MMSE /30 at baseline** **Mean (SD)**	**Age, years at first MRI** **Mean (SD)**	**Annualised loss of total (left plus right hippocampi)** **Mean % year $^{-1}$ (SD)**	**Interval (years unless specified)** **Mean (SD)**
(Jack *et al.*, 2003)	Means and SDs given by personal communication	Jack (trial)	AD	192	*20.8 (4.1)*	*72.8 (7.7)*	*5.5 (3.3)*	*1.0 (0.1)*
(Fox *et al.*, 2005)	Yes	Fox (trial)	AD - placebo	57	20.2 (3.5)	70.7 (8.2)	*3.16 (3.51)*	10.9 (1.1) months
(Wang *et al.*, 2003)	Yes	Csernansky	C	26	*29.0 (1.3) <13>*	73 (7.0)	*2.3 (1.9)*	2.2 (0.53)
			DAT	18	*25.7 (3.9) <12>*	74 (4.4)	*5.1 (2.8)*	2.0 (0.37)

Study Author, Year	Checked with author	Research Group	Subject groups	n	MMSE /30 at baseline Mean (SD)	Age, years at first MRI Mean (SD)	Annualised loss of total (left plus right hippocampi) Mean % year $^{-1}$ (SD)	Interval (years unless specified) Mean (SD)
(Kaye *et al.*, 2005)	Yes	Kaye	C	88	28.3 (1.5)	83.0 (7.0)	2.2 (6.0)	2.04 (1.42)
			Mild	27	21.7 (4.5)	76.1 (7.1)	2.9 (7.8)	
			Moderate	17	16.8 (6.7)	75.1 (6.2)	3.2 (6.8)	

Key

AD Alzheimer's disease
C control
DAT dementia of Alzheimer type
! percentage decline over scanning interval rather than rate
<> number of subjects on which information has been obtained
italicised = information from personal communication

Table 12.2 Quality assessment of studies included in the meta-analysis.

	(Jack *et al.*, 2004)	**(Du *et al.*, 2004)**	**(Hashimoto *et al.*, 2005)**	**(Barnes *et al.*, 2005)**	**(Thompson *et al.*, 2004a)**	**(Jack *et al.*, 2003)**	**(Fox *et al.*, 2005)**	**(Wang *et al.*, 2003)**	**(Kaye *et al.*, 2005)**
Cohort **Population** **Case/control** **Case series**	Case / control	Control / control	Case series	Case / control	Case / control	Case series	Case series	Case / control	Case / control
AD Diagnostic criteria **(imaging used in criteria)**	NINCDS – ADRDA General committee consensus	NINCDS – ADRDA (Imaging used to exclude other neuropathol.)	NINCDS – ADRDA MMSE >15	NINCDS – ADRDA Imaging compatible with AD	NINCDS – ADRDA	Mild-moderate AD > 50 years old	NINCDS – ADRDA Imaging supporting AD	DAT = CDR of 0.5	NINCDS – ADRDA Mixture of community and clinical cohorts

	(Jack *et al.*, 2004)	(Du *et al.*, 2004)	(Hashimoto *et al.*, 2005)	(Barnes *et al.*, 2005)	(Thompson *et al.*, 2004a)	(Jack *et al.*, 2003)	(Fox *et al.*, 2005)	(Wang *et al.*, 2003)	(Kaye *et al.*, 2005)
AD Exclusion criteria	Symptoms unrelated to AD	Other neuropathol.	Other neuropathol. / mental disorders / congnitive-affecting treatments	N/R	White matter lesions > 3mm on T2 images Substance abuse Depression	Vascular disease Non-AD disorders Major and/or unstable conditions such as seizure, PD or tumour	Significant other neurological diseases. Medications which affect cognition. Unstable regimen of drugs which may alter cognition	Other confounding neuropsychol.	N/R
APOE genotype (% E4)	C=40 Slow=55 Fast=56	N/R	AD-C=65 AD treated=54	AD=54	N/R	N/R	AD=55	N/R	C=22 Mild AD=59 Moderate =71

	(Jack *et al.*, 2004)	(Du *et al.*, 2004)	(Hashimoto *et al.*, 2005)	(Barnes *et al.*, 2005)	(Thompson *et al.*, 2004a)	(Jack *et al.*, 2003)	(Fox *et al.*, 2005)	(Wang *et al.*, 2003)	(Kaye *et al.*, 2005)
Gender (% male)	C=42 Slow=39 Fast=42	C=44 AD=40	AD-C=22 AD-T=24	C=52 AD=41	C=50 AD=50	AD-C=40 AD-T=41	AD=40	C=46 AD=61	C=48 Mild=52 Moderate=71
Treatments	N/R	N/R	Donepezil (AD-T). AD-C Vit E etc allowed	N/R	N/R	N/R	Donepezil 60% Galantamine 5% Rivistimine 21% HRT 21% VitE 32%	N/R	N/R
Drop out if RCT	N/A	N/A	N/A	N/A	N/A	N/R	15 due to no post-baseline scan	N/A	N/A

	(Jack *et al.*, 2004)	(Du *et al.*, 2004)	(Hashimoto *et al.*, 2005)	(Barnes *et al.*, 2005)	(Thompson *et al.*, 2004a)	(Jack *et al.*, 2003)	(Fox *et al.*, 2005)	(Wang *et al.*, 2003)	(Kaye *et al.*, 2005)
Assessment of controls	Normal neurological exam Community dwelling	MRI excluding neuropathol.	N/A	N/R	N/R	N/A	N/A	CDR 0.0	Yearly assessment of neuropsychol., physical and medical exam
Exclusion criteria for controls	No active neurological disorder Psychoactive medication	Clinical history of alcoholism, psychiatric illness, epilepsy, hypertension, diabetes, major heart disease or head trauma (Du *et al.*, 2003).	N/A	N/R	Same as AD	N/A	N/A	N/R	N/R
Magnet strength (Tesla)	1.5	1.5	1.5	1.5	2	0.5-1.5	1.5	1.5	1.5

	(Jack *et al.*, 2004)	(Du *et al.*, 2004)	(Hashimoto *et al.*, 2005)	(Barnes *et al.*, 2005)	(Thompson *et al.*, 2004a)	(Jack *et al.*, 2003)	(Fox *et al.*, 2005)	(Wang *et al.*, 2003)	(Kaye *et al.*, 2005)
Sites	N/R	1	1	N/R	1	38	17	1	1
Protocol (thickness (mm))	1.6	1.4 (from other publication)	1.5	1.5	N/R	1.6	1.5-1.8	1	4 (detailed in (Kaye *et al.*, 1997)
Measurement method	Manual	Semi-automated	Manual	Manual	Manual	Manual	Manual	Semi-automated	Manual
Number of raters	N/R	1	1	2	1	1	3	1 rater for template as in (Haller *et al.*, 1997) N/R for landmark placement	N/R

	(Jack *et al.*, 2004)	(Du *et al.*, 2004)	(Hashimoto *et al.*, 2005)	(Barnes *et al.*, 2005)	(Thompson *et al.*, 2004a)	(Jack *et al.*, 2003)	(Fox *et al.*, 2005)	(Wang *et al.*, 2003)	(Kaye *et al.*, 2005)
Registered scans or TIV corrected	Registered	N/R	N/R	Registered	Registered	Registered as referenced in (Jack *et al.*, 1998)	Registered	N/R	N/R
Whole Hc measured	No (small part of tail excluded)	No (small part of tail excluded) as in (Hsu *et al.* 2002)	No (white matter excluded)	No (small part of tail excluded)	No (small part of tail excluded)	No (small part of tail excluded)	No (small part of tail excluded)	No (exclusion of white matter of alveus and fimbria some of tail excluded)	No. Body of hippocampus only as in (Kaye *et al.* 1997)
Reliability **On volume unless specified**	ICC 0.91 On rate not volume	ICC 0.94	ICC 0.98 CV 2.7%	Mean absolute difference as a percentage 3%	ICC 0.99	CV=0.28	N/R	ICC 0.93 On volumes derived from scans 1 month apart.	ICC 0.90
Blinding of raters to diagnosis	Yes	Visually checked blinded to diagnosis	Yes	Yes	Yes	Yes	Yes	N/R	N/R

	(Jack *et al.*, 2004)	(Du *et al.*, 2004)	(Hashimoto *et al.*, 2005)	(Barnes *et al.*, 2005)	(Thompson *et al.*, 2004a)	(Jack *et al.*, 2003)	(Fox *et al.*, 2005)	(Wang *et al.*, 2003)	(Kaye *et al.*, 2005)
Blinding of raters to order of scans	Yes	N/R	Yes	Yes	Yes	Yes	Yes	N/R	N/R

Key
AD Alzheimer's disease
C control
DAT Dementia of Alzheimer Type
ICC Intraclass Correlation Coefficient
CV coefficient of variation
AD-C control group of AD patients
AD-T treated AD patients
N/A not applicable
N/R not reported
CDR Clinical dementia rating
MMSE mini mental state examination
neuropathol. neuropathologies
neuropsychol. neuropsychologies

Table 12.3 Studies excluded from the meta-analysis.

Study	Reason for exclusion	Group	Subject groups	N	Mean (SD) MMSE	Mean (SD) age in years	Rate of atrophy	Mean (SD) scanning interval in years unless specified
(Mungas *et al.*, 2005)	VaD included	Weiner DeCarli	C (CDR 0.0)	58	29.0 (1.4)	74.1 (6.7)	1.1 (1.4)	4.0 (1.1-6.9)
			Dementia : AD and VaD (CDR ≥ 1.0)	11	23.6 (3.4)	76.5 (9.7)	2.9 (2.0)	2.9 (1.1-5.1)
(Cardenas *et al.*, 2003)	VaD included	Weiner	C	16	29.1 (1.0)	76 (5)	1.8 (0.8)	2.6 (1)
			Dementia	7 (5 AD) 2 (AD + VaD)	24.6 (2.6)	76 (8)	5.4 (2.8)	2.6 (1)
(Moffat *et al.*, 2000)	Controls only	Resnick	C ε4 +	13	27.9 (2.5)	68.5 (5.9)	2.86 χ	2.7 (0.6)
			C ε4 -	13	29.1 (1.1)	69.7 (6.8)	0.85 χ	2.6 (0.8)
(Jack *et al.*, 1998)	Overlap with Jack *et al.* 2000 paper	Jack	C	24	28.79 (1.28)	81.04 (3.78)	1.55 (1.38)	1.96 (0.75)
			AD	24	20.74 (4.60)	80.42 (4.02)	3.98 (1.92)	1.89 (0.68)

Study	Reason for exclusion	Group	Subject groups		N	Mean (SD) MMSE	Mean (SD) age in years	Rate of atrophy		Mean (SD) scanning interval in years unless specified
(Mori *et al.*, 2002)	Overlap with Hashimoto *et al.* 2005	Hashimoto	AD ε4 +		38	19.2 (5.1)	72.4 (5.5)	9.76 (4.27)		378 (18) days
			AD ε4 -		17	19.1 (6.3)	71.1 (6.4)	6.99 (4.24)		372 (20) days
(Silbert *et al.*, 2003)	Overlap with Kaye *et al.* 2005	Kaye	C ⇑		8	28.4 (1.4)	86.6 (8.5)	2.6 (3.2)		4.1 (2.2)
			Demented⇑		20	22.8 (7.4)	81.4 (8.5)	3.1 (5.7)		
(Krishnan *et al.*, 2003)	Short interval	Rogers	Placebo AD		33‡	19.0 (4.6) ‡	72.4 (10.1) ‡	8.2 (9.9)!		24 wks
			Treated AD		34‡	19.5 (4.8) ‡	74.4 (7.0) ‡	-0.4 (2.9)!		24 wks or endpoint
(Jack *et al.*, 2000)	Overlap of subjects with Jack *et al.* 2004	Jack	C total	Stab	58 = 48	28 (1.6)	80.4 (6.4)	1.9 (1.1)	1.7 (0.9)	3.0 (0.5)
				conv	+ 10	28 (1.7)	82.3 (5.8)	total	2.8 (1.7)	3.3 (0.4)
			AD		28	22 (4.3)	73.8 (11.3)	3.5 (1.8)		2.9 (0.5)

Study	Reason for exclusion	Group	Subject groups	N	Mean (SD) MMSE	Mean (SD) age in years	Rate of atrophy	Mean (SD) scanning interval in years unless specified
(Kaye *et al.*, 2005)	Subgroup excluded (MMSE > 26)	Kaye	v.mild	23	26.2 (2.0)	83.0 (7.8)	4.3 (7.4)	2.04 (1.42)
(Hampel *et al.*, 2005)	Mean rates not published. Methodological problems and errors in paper.	Hampel	AD	22	23.1 (4.0)	67.8 (7.9)	15.6 (10.5)	18.4 (9.4) months
(Jack *et al.*, 2004)	Subgroup excluded (MMSE > 26)	Jack	C conv	15	28.4 (1.4) <12>	80.2 (4.6)	3.1 (1.6)	3.6 (0.9)
(Laakso *et al.*, 2000b)	Cannot provide original data annualised	Soininen	C	8	28.4 (1.6)	69 (8)	1.2 (5.0)†	36 (8.9) months
			AD	27	22.0 (3.6)	70 (5)	2.4 (6.7)†	

Key

χ No SD given on the % loss. Mean (SD) rate initially quoted as 71.2 (19.8) ε4+ and 20.1 (16.5) ε4- mm^3 loss per year.
⇧Only 28 subjects had more than one scan and therefore summary statistics are slightly incorrect.
^ median and min-max
* Standard error rather than standard deviation
⌂ No SD given on the % loss. Percentage loss expressed over the two year interval as (controls: left 4.0, right 5.5, DAT: left 8.3, 10.2 % loss)
‡ summary statistics given for a wider group than on which hippocampal changes were calculated.
† calculated rates of atrophy (loss quoted in publication over 3 year period actual rates over 3 years 3.6 (15.1) for controls and 7.2 (20.1) % of hippocampus.)
VaD = vascular dementia

Data synthesis and heterogeneity

Figure 12.1 is a Forest plot demonstrating the results of the AD subjects from the individual studies included in the meta-analysis (n=591). The estimate of the overall mean atrophy rate in these groups from a random-effects meta-analysis was 4.71% per year (95% CI 3.90, 5.52). The I^2 squared statistic from this analysis was 78.5%, suggesting that most of the variability observed between the studies is true heterogeneity as opposed to sampling variability. Figure 12.2 is a Forest plot of the matched control groups (n=243). The estimated overall mean atrophy rate in these groups from a random-effects meta-analysis was 1.52% per year (95% CI 0.89, 2.15). The I^2 squared statistic was 59%, suggesting just over half of the observed variability was due to genuine heterogeneity, as opposed to sampling variability. Figure 12.3 shows the differences between the control and AD groups in those studies where both groups were studied (six studies in total). Using data from these studies, and combining AD subgroups within studies, using a random effects meta-analysis the estimate of the mean difference in atrophy rates between controls and AD subjects was 3.35% per year (95% CI 1.71 to 4.99). The I^2 estimate for this AD-control difference was 74.8%, again suggesting that most of the observed variability in differences is due to true heterogeneity.

Publication bias

Figure 12.4 and Figure 12.5 show funnel plots of AD and AD-control differences respectively against the standard error of the estimate. Although few conclusions can be drawn given the small number of studies, there was no suggestion of publication bias.

Subgroup analysis and meta-regression

Plots of the study/subgroup mean atrophy rates against age, MMSE and inter-scan interval are shown in Figure 12.6, to Figure 12.8 respectively. There were no apparent relationships between these variables and mean atrophy rate. The corresponding meta-regression models showed that there was no evidence that these variables were associated with mean rates of atrophy in the AD groups ($p>0.05$).

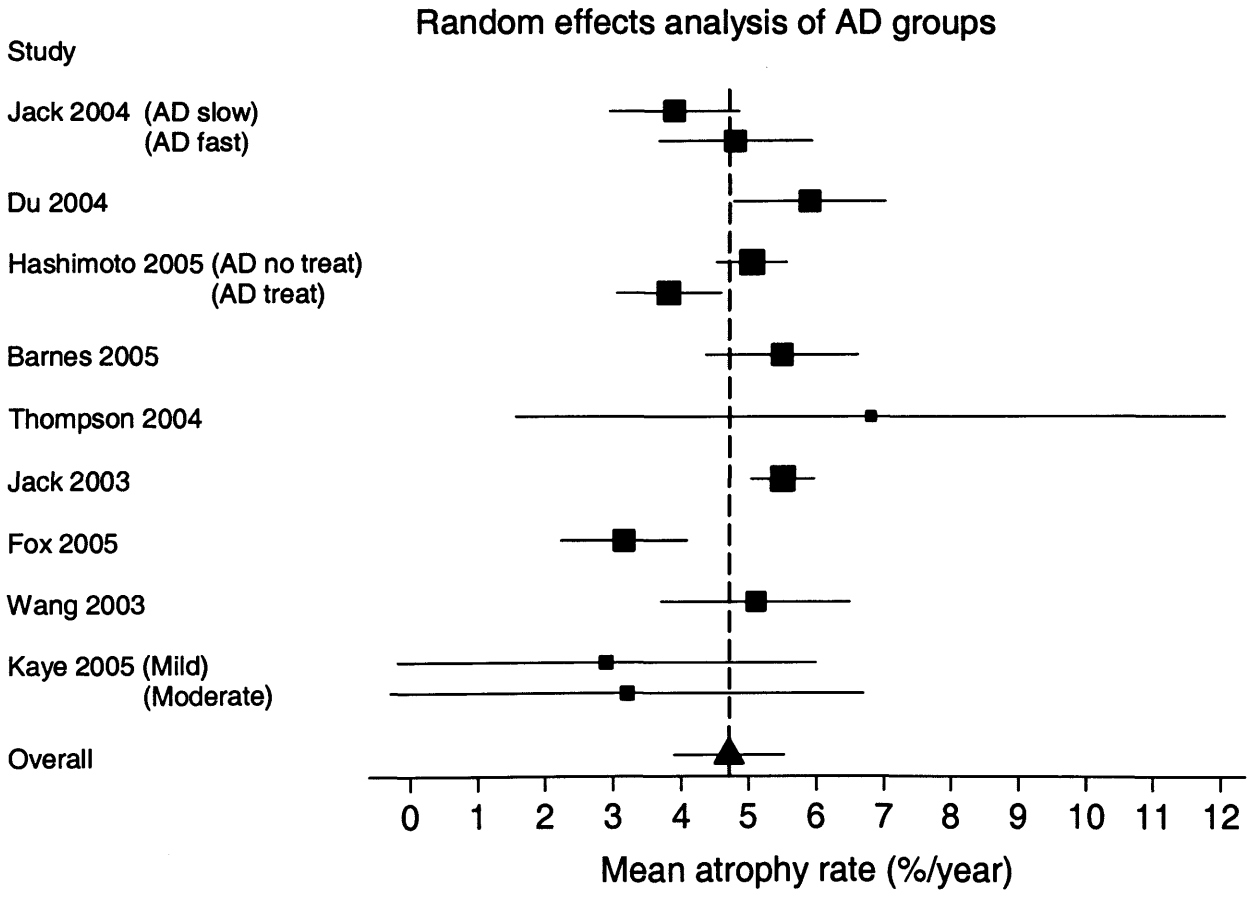


Figure 12.1 Forest plot of rates of atrophy in the hippocampus in AD subjects allowing for sub-groups of the same study to be entered separately. The size of the squares are proportional to 1/(variance(study i) + between-study variance). The solid lines represent 95% CIs.

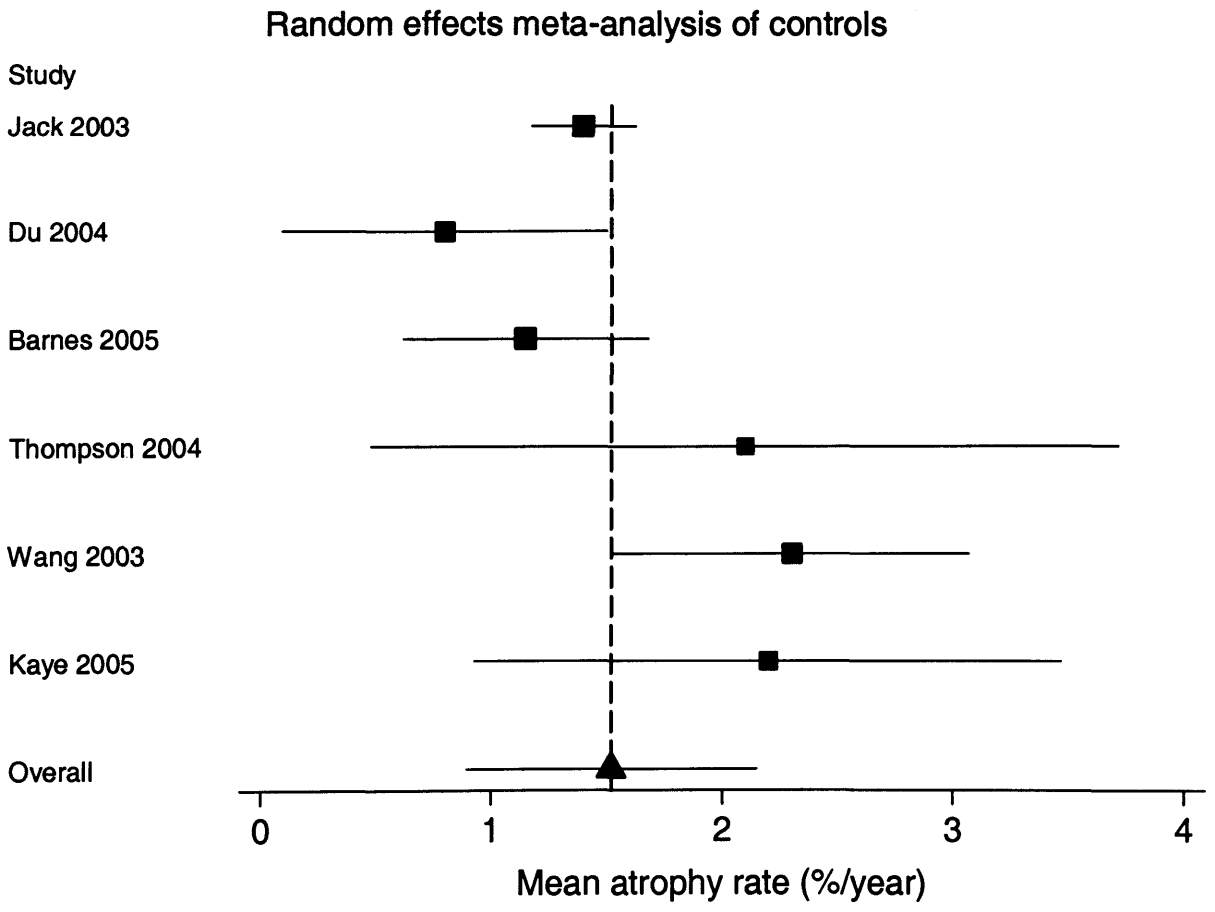


Figure 12.2 Forest plot of rates of atrophy in the hippocampus in matched control subjects.
The size of the squares are proportional to 1/(variance(study i) + between-study variance). The solid lines represent 95% CIs.

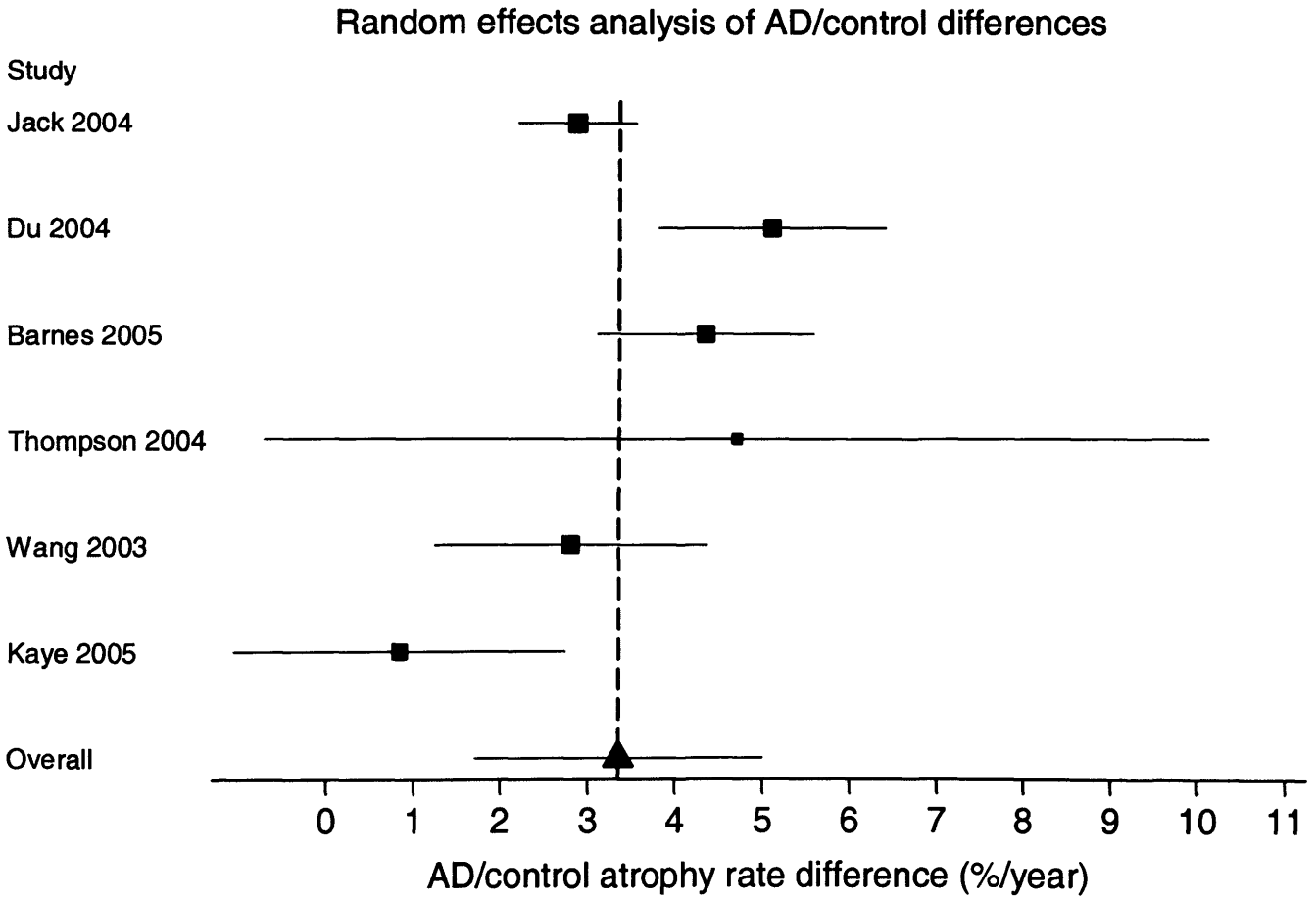


Figure 12.3 Forest plot showing the difference between AD and control subjects in studies where both were reported. The size of the squares are proportional to 1/(variance(study i) + between-study variance). The solid lines represent 95% CIs.

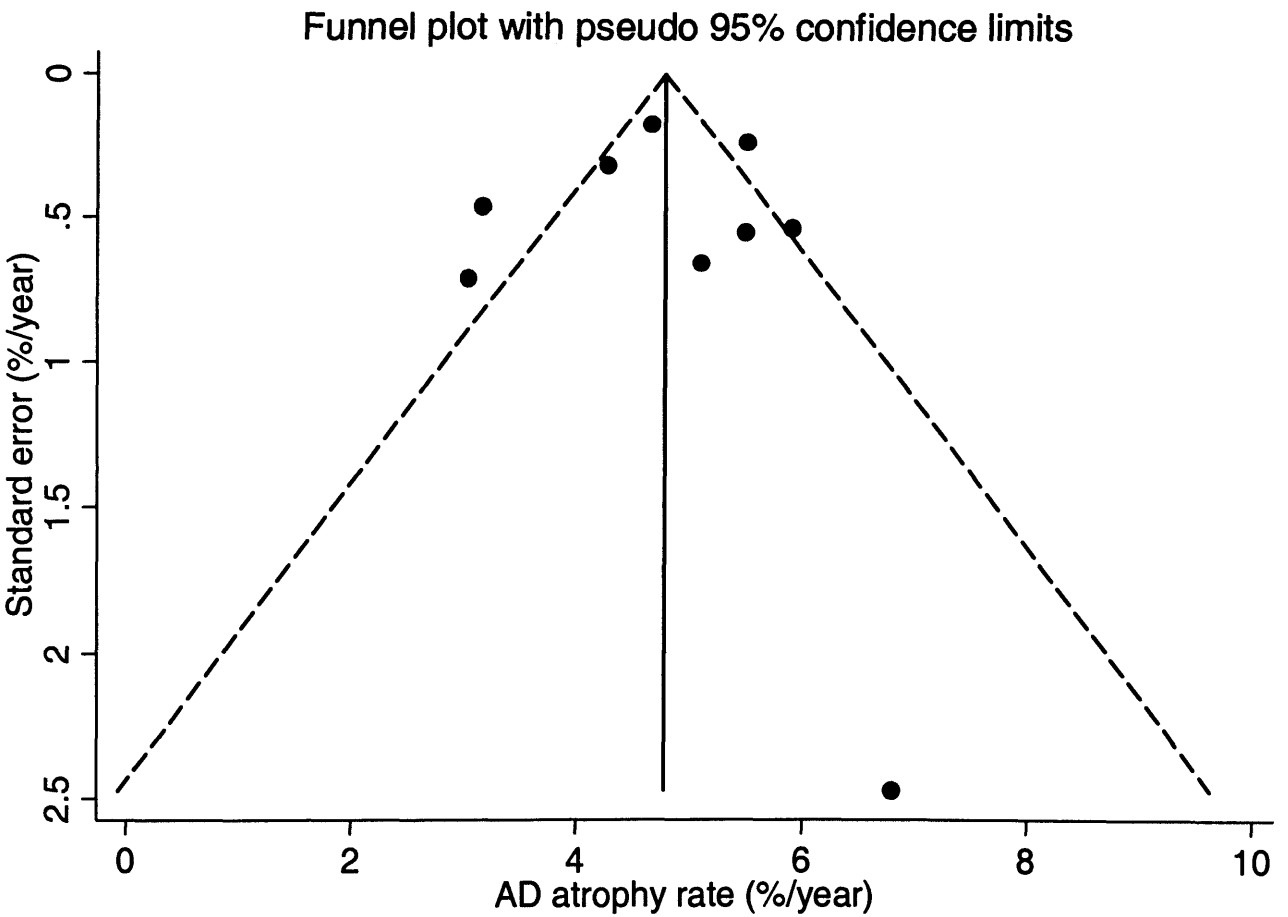


Figure 12.4 Funnel plot of AD rates of atrophy with pseudo 95% confidence limits.

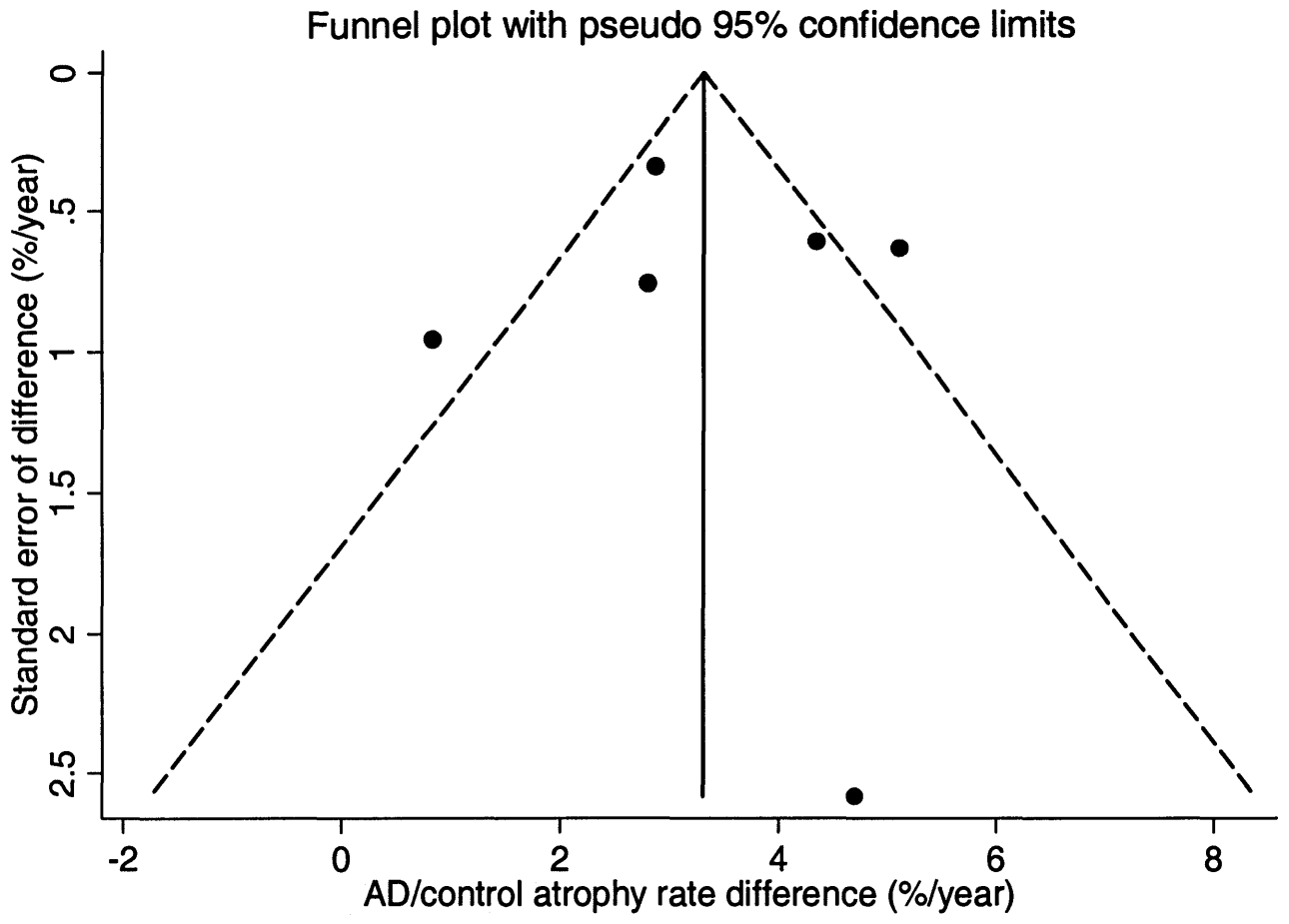


Figure 12.5 Funnel plot of AD-control differences with pseudo-95% confidence limits.

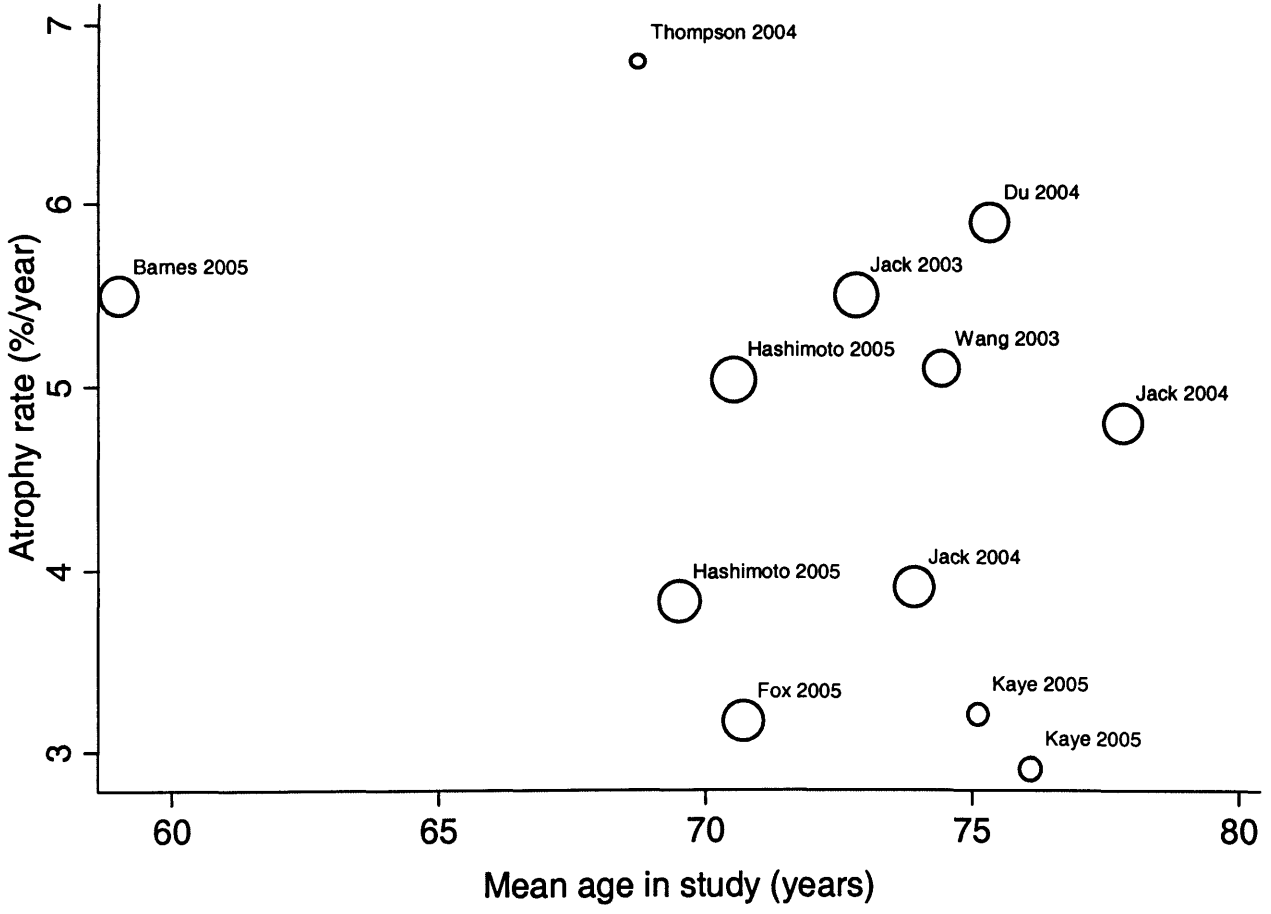


Figure 12.6 Meta-regression of atrophy rate and mean age.
The size of the circles are proportional to 1/(variance(study i) + between-study variance).

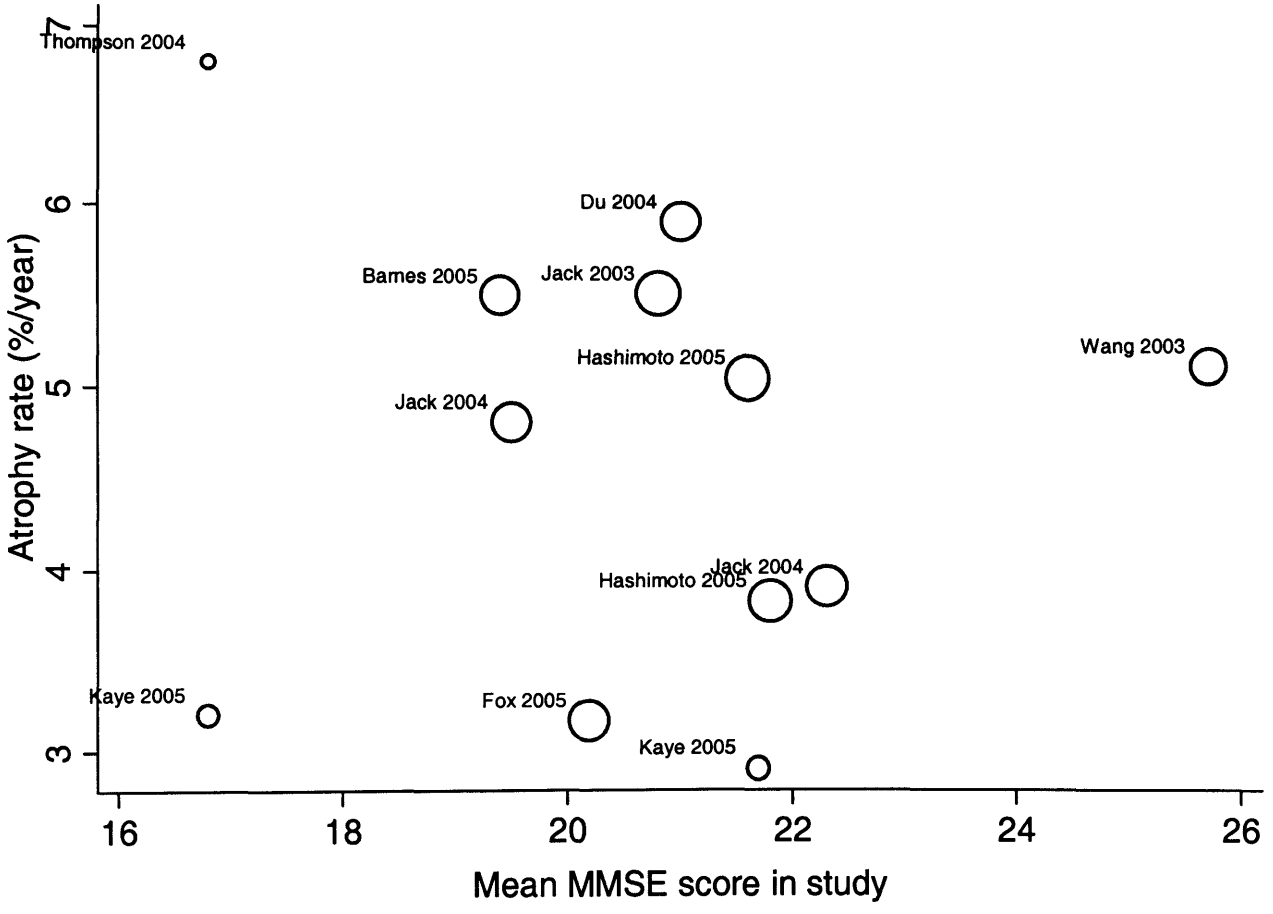


Figure 12.7 Meta-regression of atrophy rate and mean MMSE score.
The size of the circles are proportional to 1/(variance(study i) + between-study variance).

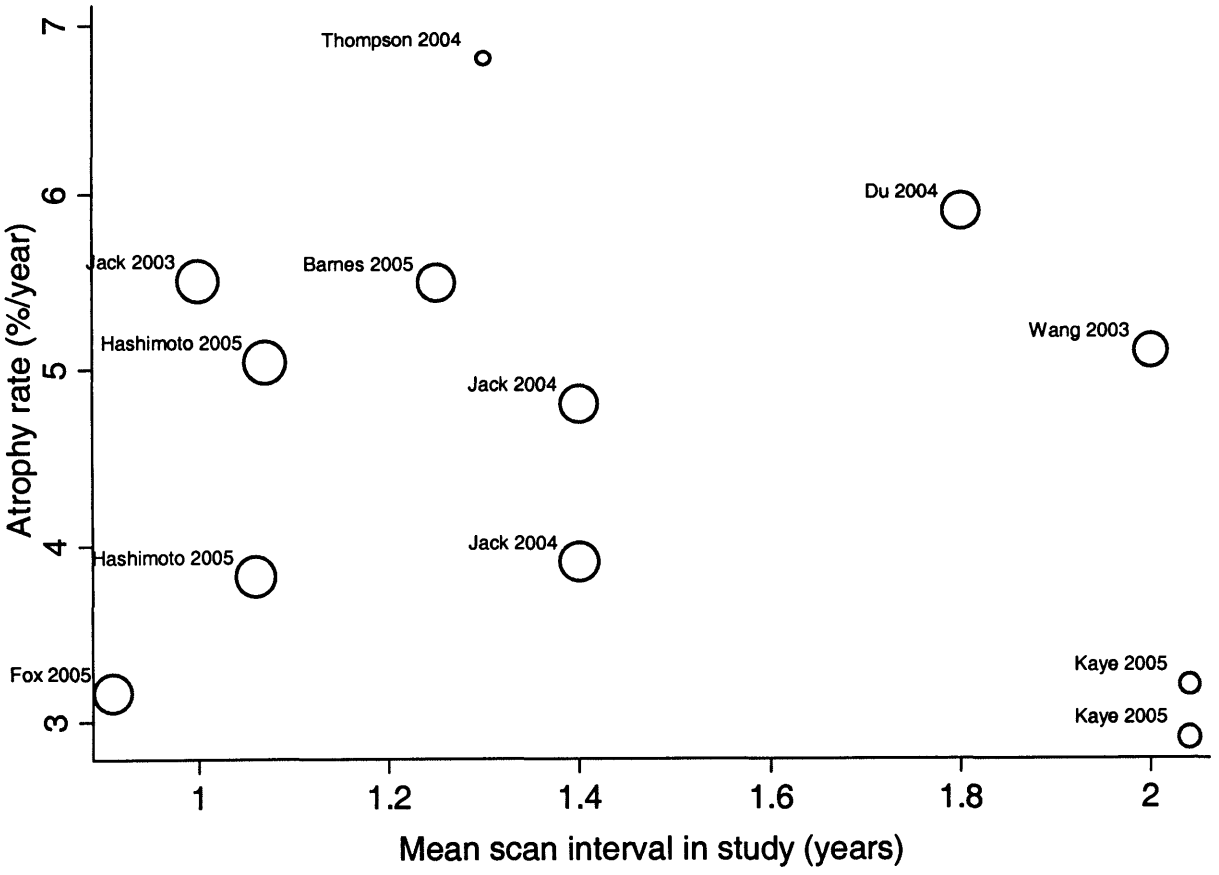


Figure 12.8 Meta-regression of atrophy rate and mean scan interval.
The size of the circles are proportional to 1/(variance(study i) + between-study variance).

12.5. Discussion

This study aimed to estimate the mean rates of hippocampal atrophy across a number of published studies. In addition, this study aimed to formally assess some aspects of the heterogeneity in these studies. Meta-analysis plays an important role in summarising results from studies owing to between-study heterogeneity. It may be that the effects of confounders such as disease severity or age are lessened when the results from all studies are pooled, making the results more applicable to the wider population.

It was found that hippocampal rates of atrophy were consistently greater in AD subjects than controls in all studies included in this meta-analysis. Synthesised mean rates of atrophy were 1.5% in the control subjects and 4.7% in AD subjects. This is consistent with the mean difference between AD and control groups being 3.4% for the subset of studies where both controls and AD subjects had been assessed and reported. Most of the heterogeneity in these studies was due to true heterogeneity rather than sampling error.

Some aspects of the heterogeneity of studies are difficult to formally assess due to the lack of reporting of certain variables (see Table 12.2). These variables include patient-related information such as other administered medications and APOE genetic status. Treatments may be particularly difficult for case-control or population-based studies to report accurately, since many of these studies will be conducted at a tertiary referral centre whereas it may be the general physician at the primary level who is responsible for both prescribing and monitoring the administration of medication.

Other scan and hippocampal delineation methods may also have an effect on the atrophy rates being determined such as scan "slice" thickness or segmentation protocol. It may be that differences in anatomical structures included in the delineation of hippocampi do not have substantial effects on the atrophy rate determined if the same protocol is used for both baseline and follow-up scans. This implies that rates of atrophy are not different in separate anatomical locations within the hippocampus (for example, rate of atrophy of the hippocampal tail is similar to the hippocampal head). Meta-analysis on the AD-control differences should be robust to these potential confounders since the hippocampal rates determined in both subject groups should have been determined in the same way.

Formal meta-regression analyses showed no significant associations between mean atrophy rates quoted in the studies included in the meta-analysis and age, MMSE and interval. However, it may be that these relationships do exist; the ability of meta-regression to detect such associations depends upon there being sufficient variability in the explanatory variable between studies. There was little variability in mean age between studies relative to within studies, perhaps explaining the lack of association. A number of studies (not included here) have examined rates of hippocampal atrophy in normal ageing and have shown that age does influence hippocampal atrophy rates in "healthy controls". There was some variability in mean MMSE score between studies. The included studies contained a wide range of mean interval between scans, and thus the data suggest no association between mean interval and mean atrophy rate.

This study has a number of limitations. Although this meta-analysis attempted to be as inclusive of as many studies as possible and there was no evidence of publication bias, there were only nine studies that were analysed owing to a number of reasons (see Table 12.3). Larger numbers of studies would allow a more precise estimate of the mean rates to be calculated and for associated meta-regression analyses to be more robust. Collating the individual patient data from these studies would enable much more precise estimation of relationships between factors such as age and disease severity with hippocampal atrophy rates. Associations using individual patient data within studies would not be confounded by study-level factors, such as scan acquisition protocols and methods to determine hippocampal atrophy rate. Also, variability both within and between studies could be used to increase the precision of estimation of these associations.

12.6. Chapter Conclusions

The hippocampal atrophy rate is on average 1.5% in normal controls with the range of the quoted mean age ages being 59-83 years. In AD subjects this rate is on average 4.7%. Age, disease severity (determined by the MMSE) and scanning interval did not appear to be associated with these mean rates within the AD groups.

13. THESIS CONCLUSIONS

13.1. Clinical findings

The main focus of this thesis was to assess rates of atrophy in AD and try and improve upon the operator-intensive nature of measures of atrophy of small areas of the brain (particularly the hippocampus) currently used for diagnosis and assessment of progression in AD. In order to assess new techniques, the "gold standard" manual-outlining method of atrophy required analysis. In order to understand the potential benefits and limitations of analysis of small regions in clinical practice, the ways in which these regions are affected by disease needed to be appreciated.

A protocol for delineation of the cingulate gyrus, an area of the brain highlighted in many studies as a region specifically affected by AD, was developed. The cingulate gyrus was shown to be smaller in AD compared with controls, particularly posteriorly. Applying this technique to longitudinal images in groups of pathologically confirmed AD, FTLD and age-matched controls, the cingulate was shown to atrophy at a greater rate in disease groups than controls, with FTLD overall showing a greater atrophy rate than AD. In both groups the atrophy rate of the cingulate exceeded that of the hippocampus, allowing the conclusion to be drawn that the cingulate is at least as affected in AD and FTLD as the hippocampus. Differences were also seen in the pattern of atrophy along the length of the hippocampus between the AD and FTLD group with relatively more posterior compared with anterior atrophy in AD and the reverse in FTLD. Rates of cingulate atrophy have not previously been reported in these groups. This study shows the use of regional analysis in the differential diagnosis of dementia.

Measures of hippocampus and amygdala volume were calculated in a relatively large group of pathologically confirmed AD and FTLD. The results from this analysis showed that both the amygdala and hippocampus were significantly affected in AD and FTLD. Further analysis showed that the hippocampal to amygdala ratio was similar in AD to controls but that the volumes of both structures were reduced. This ratio was altered in FTLD with the amygdala being disproportionately smaller in FTLD. This is consistent with the anterior/posterior gradient of atrophy described in the cingulate in FTLD. Overall, FTLD subjects had lower volumes than AD subjects although there was difficulty assessing and comparing the severity of these two diseases. Dividing the

FTLD group into its subtypes, this study showed that the differential pattern of atrophy between the hippocampus and amygdala on the left and right sides differed according to clinical rather than pathological subtype. Semantic dementia patients showed greater asymmetry whilst fvFTD subjects showed greater amygdala involvement. Again, this type of analysis has implications for the differential diagnosis of dementia.

Many studies, including those described as part of this thesis, have evaluated rates of atrophy of the hippocampus with inter-scan intervals at one year or greater. Chapter 5 showed that hippocampal change could be used to separate AD subjects from controls at intervals of less than one year, specifically six months. Rates of atrophy of the hippocampus were greater in AD subjects than controls and were comparable at a group level (although less precise) than the corresponding rates calculated using a 12 month interval. This is an interesting finding since longitudinal measures of atrophy of areas such the hippocampus may be useful in clinical trials, where the cost of the trial may be substantially reduced by shorter follow-up intervals. This gain is offset by the greater sample size (albeit non-significantly greater as reported in Chapter 5) needed as a result of the higher variance in measured rates of atrophy.

In one further study the asymmetry of the hippocampus was evaluated as there was some evidence from the literature that a change in the normal right > left asymmetry could be a potential maker of AD. Although it was found that there was some evidence of a reduction in this normal asymmetry resulting from greater right than left rates of atrophy in AD, this was by no means a consistent finding, and it was therefore concluded that this was not a useful marker of AD. As part of this same study, the effect of APOE status (ε4 positive carriers vs. ε4 negative carriers) on both hippocampal volume and rate of atrophy was assessed. Only a subgroup of the AD subjects had this information available and it was not possible to show that APOE ε4 status had a significant effect on volume or rate of atrophy of the hippocampus. However, this subgroup was small and a larger average atrophy rate was seen in the ε4 positive carriers. Were this analysis performed in a larger group or using a method which reduced manual errors such a study may detect this to be a significant effect.

A study of the hippocampal region in familial AD subjects revealed that longitudinal changes in this structure could be measured approximately five and a half years prior to clinical diagnosis of AD, whereas whole brain change could be detected approximately

three and a half years before diagnosis. This finding highlights the importance of the development and improvement of measures of such change in order that early disease may be detected and "disease modifying" medication administered as early as possible should it become available.

Finally, a meta-analysis of hippocampal atrophy rates was performed to synthesise a rate of atrophy from those rates already published. This rate was calculated as 4.7% per year in AD subjects, with age, severity of disease (measured using MMSE) and scanning interval having no significant effect on the rates of atrophy in the AD subjects.

13.2. Image analysis developments

Segmentation methodology for the cingulate gyrus was developed and evaluated as part of this thesis. This work showed that areas different in morphology to those often segmented in AD (hippocampus, amygdala and entorhinal cortex) could be delineated reliably. Moreover, these regions could be subdivided reproducibly. Chapter 7 in this thesis showed that subdivision could also be automated by the application of Bezier curves, a technique widely used in computer graphics. Development of such measures provides useful information with little operator intervention and may help to reveal differential patterns of atrophy between subject groups. In addition, this type of subdivision may be applicable to other structures in the brain.

The majority of the technical work within this thesis focused on the hippocampus, specifically on building upon current techniques and methodology to enable automatic serial analysis of the structure to aid diagnosis and disease tracking in AD. In all studies, comparisons of new techniques or developments were made with manual measures.

First application of the BSI to the hippocampus was described. This was found to be effective at reducing the manual error associated with serial hippocampal segmentation and as a result to have diagnostic utility. The HBSI was also applied to a large trial dataset (17 sites internationally) and was shown to give results which were similar to manual measures in the placebo arm and revealed a significant difference between that arm of the trial and the treatment arm.

Secondly, another technique, fluid-propagation of hippocampi and Jacobian calculation, first described in a separate study, was applied to a larger group of subjects. Advances to the procedure were made including an extra local rigid alignment of hippocampi prior to fluid-registration and development of exit criteria allowing all hippocampal pairs to attain the same level of matching. These fluid-based techniques also showed a reduction in the human error associated with longitudinal hippocampal demarcation. Both HBSI and fluid-techniques may be particularly useful as they reduce operator time by eliminating the need for hippocampal segmentation of one of the pair of scans.

Thirdly, the hippocampal BSI technique was completely automated by the application of an affine alignment of template hippocampus to the baseline of every image pair. Evaluations were made to establish whether improvements could be made to this initial step. Simple operations of region intensity thresholding and conditional dilations improved the accuracy of this initial segmentation. However, the BSI measure was robust enough that this did not make substantial improvements to the measure of change as assessed by similarity to the standard HBSI measure (calculated using the manually segmented baseline region). Averaging a number of regions from half of the cohort to create a population-based template also slightly improved the baseline template-based segmentation at some probability thresholds, but did not improve the similarity of the automated HBSI measure to the standard HBSI measure.

Finally, the location of differences in the hippocampus was assessed using a voxel by voxel test of hippocampal masks, affine-registered whole brain – whole brain followed by rigid-registered hippocampus – hippocampus to a single-person cohort template image. These assessments showed that regions such as the superior of the hippocampus were greatly reduced in AD compared with controls at both baseline and repeat imaging. However, this study also revealed the AD subjects had more voxels labelled as hippocampus in the inferior of the region, potentially caused by differential registration errors in AD and control groups, or differential shape changes of the hippocampus which were not accommodated by the final rigid registration step. This study highlights the problems associated with making judgments of location-specific atrophy caused by a disease using inter-subject registration with voxel by voxel significance tests. Further work is needed to ascertain whether non-linear registration can improve the hippocampal match without removing the location-specific atrophy which the experiment is designed to detect.

13.3. Conclusions

This thesis shows that measures of small regions of the brain can be made reliably and can reveal distinctive patterns which may aid the diagnosis and tracking of Alzheimer's disease and other dementias. Such measures can be made with less operator involvement with application of registration-based methodologies, reducing both time and associated human errors.

PUBLICATIONS

This section documents published book chapters and peer-review papers based on the results in this thesis. I would like to thank everyone who was involved in generating and analysing these results and their contributions are stated below.

Chapter 1

Barnes J, Archer H and Fox NC

Imaging in Alzheimer's Disease

Book chapter for: Research Progress in Alzheimer's Disease and Dementia (*In press*)

Clinical contributions to this were made by Hilary Archer and this was drafted by Nick Fox, Hilary Archer and me.

Chapter 4

Jones BF, **Barnes J**, Uylings HB, Fox NC, Frost C, Witter MP, Scheltens P.

Differential Regional Atrophy of the Cingulate Gyrus in Alzheimer Disease: A Volumetric MRI Study.

Cereb. Cortex. 2006; 16 (12): 1701-1708

Measurements of the cingulate and its subdivisions were made by Bethany Jones and myself. Bethany performed measurements of the RS alone. Analysis was performed by Chris Frost. Drafting of the publication was performed by Bethany Jones and me. Supervision and guidance was given by Philip Scheltens and Nick Fox. Harry Uylings and M Witter provided supportive anatomical information.

and

Barnes J, Whitwell JL, Frost C, Josephs KA, Rossor MN, Fox NC

Measurements of the amygdala and hippocampus in pathologically confirmed Alzheimer's disease and frontotemporal lobar degeneration.

Archives of Neurology. 2006; 63 (10): 1434-1439

Post-mortem reports were examined by Keith Josephs. Measurements were performed by Jenny Whitwell (amygdala) and me (hippocampus). Chris Frost and I performed the analysis. The manuscript was drafted by Jenny Whitwell and me. Supervision was given by Martin Rossor and Nick Fox.

Chapter 5

Barnes J, Godbolt AK, Frost C, Boyes RG, Jones BF, Scahill RI, Rossor MN, Fox NC.
Atrophy rates of the cingulate gyrus and hippocampus in AD and FTLD.
Neurobiol Aging 2007; 28 (1): 20-28.
Alison Godbolt assessed patient information and I performed all the measurements. Chris Frost and I performed the analysis. Richard Boyes helped with some technical details. I drafted the manuscript. Rachael Scahill, Martin Rossor and Nick Fox supervised the study.

and

Barnes J, Scahill RI, Frost C, Schott JM, Rossor MN, Fox NC.
Increased hippocampal atrophy rates in AD over six months using serial MR imaging
Neurobiology of Aging. 2007 (In Press)
I performed all the measurements and the analysis under the supervision of Chris Frost. I drafted the manuscript under the supervision of Nick Fox, Martin Rossor and Rachael Scahill. Jon Schott collected all the scans.

Chapter 6

Barnes J, Scahill RI, Schott JM, Frost C, Rossor MN, Fox NC.
Does Alzheimer's disease affect hippocampal asymmetry? Evidence from a cross-sectional and longitudinal volumetric MRI study.
Dement Geriatr Cogn Disord. 2005;19(5-6):338-44.
Rachael Scahill and I performed all the measurements. Chris Frost and I performed the analysis. Jonathan Schott helped with drafting the manuscript. Martin Rossor and Nick Fox supervised the project.

and

Ridha BH, **Barnes J**, Bartlett JW, Godbolt AG, Pepple T, Rossor MN, Fox NC.
Tracking atrophy progression in familial Alzheimer's disease: a serial MRI study.
Lancet Neurology 2006; 5 (10): 828-834.
Basil Ridha and I evaluated the scans that were used and performed all the hippocampal measurements. Tracey Pepple calculated BSI measures. Jonathan Bartlett and Basil Ridha performed the analyses. Alison Godbolt evaluated patients and assessed their clinical severities by date. Basil Ridha, Jonathan Bartlett and I drafted the manuscript under the supervision of Nick Fox and Martin Rossor.

Chapter 7

Boyes RG, **Barnes J** and Fox NC

MIUA conference proceedings 2004

Partitioning the Cingulate Gyrus using Bezier curves

Richard Boyes wrote the code for subdividing the cingulate using Bezier curves. I made all the measurements and performed the analysis. Richard and I drafted the short conference paper. Nick Fox supervised the project.

Chapter 8

Barnes J, Scahill RI, Boyes RG, Frost C, Lewis EB, Rossor CL, Rossor MN, Fox NC. Differentiating AD from aging using semiautomated measurement of hippocampal atrophy rates.

Neuroimage. 2004 Oct;23(2):574-81.

Rachael Scahill and I performed all the manual measurements. Richard Boyes and Emma Lewis helped with technical aspects with application of the BSI to the hippocampus. Charlotte Rossor provided some extra measurements to aid our understanding of the application of the BSI to the hippocampus. I performed the analysis under the supervision of Chris Frost. I drafted the manuscript. Martin Rossor and Nick Fox provided supervision for the project.

and

Barnes J, Lewis EB, Scahill RI, Bartlett JW, Frost C, Schott JM, Rossor MN, Fox NC. Automated measurement of hippocampal atrophy using fluid-registered serial MRI in AD and controls.

JCAT: 2006 (In Press)

Rachael Scahill and I performed all the manual measurements. Emma Lewis and I analysed output from the fluid code. Jonathan Bartlett and I performed the statistical analysis. I drafted the manuscript with input from Emma Lewis, Nick Fox and Jonathan Bartlett. Martin Rossor, Nick Fox and Rachael Scahill provided supervision for the project.

Chapter 9

Barnes J, Boyes RG, Lewis EB, Schott JM, Frost C, Scahill RI, Fox NC

Automatic calculation of hippocampal atrophy rates using a hippocampal template and the boundary shift integral

Neurobiology of Aging 2006 (In press but available online)

I performed all segmentations of the hippocampi and developed the registration algorithm described in the publication. Richard Boyes and Emma Lewis provided technical support. Jonathan Schott recruited and assessed all the subjects in the MIRIAD study. Chris Frost provided statistical analysis advice and Chris and I performed the analyses. Rachael Scahill and Nick Fox provided supervision and gave advice regarding both the methods and the drafting. The majority of the drafting was done by myself but all authors had input into this.

ACKNOWLEDGEMENTS

I am extremely thankful for the time given by all the patients and their partners and carers to take part in research studies at the Dementia Research Centre. I would also like to extend my thanks to those people who have helped scan these subjects, at St Mary's Hospital, the National Hospital for Neurology and Neurosurgery (on the main site and Queen Square Imaging Centre) and the Institute of Neurology, particularly "Evil" Dave MacManus. All of people involved have provided the quality scans necessary for the analysis performed in this thesis.

I would like to give special thanks to the Medical Research Council for their funding of my position within the Dementia Research Centre, without their support I would not have been able to complete this thesis. There are many people to thank at the Dementia Research Centre for their support in the production of this thesis. Thanks to those who directly helped with the work which formed this thesis: Rachael Scahill, Richard Boyes, and Emma Lewis in particular, but also the contributions of Jenny Whitwell, Jo Foster, Tracey Pepple, Charlotte Rossor, Tom Yeatman and Jon Schott. Thanks to Dave Cash from CMIC for his thoughtful contributions. I very much appreciate the supervision from Nick Fox, Rachael Scahill and Martin Rossor without whose expert guidance the whole PhD process would have been very difficult. I would like to thank Chris Frost and Jonathan Bartlett for all their help and supervision with statistics. Thanks also to the support staff, namely Moira Young and Gillian Barley for arranging supervisions and made sure I submitted all the necessary paperwork in time. Thanks also to other people at work with whom I have discussed various issues, whether or not they are related to work: Imaging Team: Shona Price, Valerie Anderson, Ali Bendriss, Musib Siddique, Caroline Sherrard, Nicola Hobbs; Psychologists: Susie Henley, Seb Crutch, Katy Randlesome and Julia Hailstone; Medics: Alison Godbolt, Basil Ridha, Jon Schott and Hilary Archer; Administrative Support: Suzie Barker, Anne Parnell and Claire Bloomfield. Thanks very much to Alex, who contributes to working life on a daily basis.

I would also like to thank members of other research groups for their help, without whom, the meta-analysis which forms Chapter 12, would not be possible: Laura van de Pol (Amsterdam), Clement Loy (Sydney, Australia), Clifford Jack (Rochester, USA), Lei Wang (Washington, USA), Jeffrey Kaye (Oregon, USA), Paul Thompson (Los Angeles, USA), Mamoru Hashimoto (Japan), and An-Tao Du (San Diego, USA).

Last but no means least thank you very much to Duncan and the furries for all their support. Thanks also to my close friends Bamit, Swatkin, Anna, Stiff and everyone else I have no more room to mention.

Appendix One: DIAGNOSTIC CRITERIA

Diagnosis was made on the basis of the criteria detailed below. Assessments were performed by clinicians involved in the Specialist Cognitive Disorders Clinic.

DEMENTIA

Criteria given by the Diagnostic and Statistical Manual of Mental Disorders (4th Edition) (DSM-IV) (American Psychiatric Association, 1994):

I. The development of multiple cognitive deficits manifested by both:

a) Memory impairment (impaired ability to learn new information or to recall previously learned information)

b) One (or more) of the following cognitive disturbances are required for diagnosis:

(i) Aphasia (language disturbance)

(ii) Apraxia (impaired ability to carry out motor activities despite intact motor function)

(iii) Agnosia (failure to recognize or identify objects despite intact sensory function)

(iv) Disturbance in executive functioning (i.e., planning, organizing, sequencing, abstracting)

II. The cognitive deficits in Criteria Ia and Ib each cause significant impairment in social or occupational functioning and represent a significant decline from a previous level of functioning

III. The deficits do not occur exclusively during the course of a delirium

Dementia if:

All criteria (I, II and III) are answered with "yes"

(Note that only one of the features listed under Ib has to be present)

ALZHEIMER'S DISEASE

Criteria given by the National Institute of Neurological and Communicative Disorders and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) (McKhann *et al.*, 1984):

PROBABLE ALZHEIMER'S DISEASE

I. The diagnosis of PROBABLE Alzheimer's disease is supported by:

a) Dementia established by clinical examination and documented by the Mini-Mental Test; Blessed Dementia Scale, or some similar examination, and confirmed by neuropsychological tests

b) Deficits in two or more areas of cognition

c) Progressive worsening of memory and other cognitive functions

d) No disturbance of consciousness

e) Onset between ages 40 and 90, most often after age 65

f) Absence of systemic disorders or other brain diseases that in and of themselves could account for the progressive deficits in memory and cognition

II. The diagnosis of PROBABLE Alzheimer's disease is supported by:

a) Progressive deterioration of specific cognitive functions such as language (aphasia), motor skills (apraxia), and perceptions (agnosia)

b) Impaired activities of daily living and altered patterns of behaviour

c) Family history of similar disorders, particularly if confirmed neuropathologically

d) Normal lumbar puncture as evaluated by standard techniques

e) Normal pattern or non-specific changes in EEG, such as increased slow-wave activity

f) Evidence of cerebral atrophy on CT with progression documented by serial observation

III. Other clinical features consistent with the diagnosis of PROBABLE Alzheimer's disease:

a) Plateaus in the course of progression of the illness

b) Associated symptoms of depression, insomnia, incontinence, delusions, illusions, hallucinations, catastrophic verbal, emotional, or physical outbursts, sexual disorders, and weight loss

c) Other neurologic abnormalities in some patients, especially with more advanced disease and including motor signs such as increased muscle tone, myoclonus, or gait disorder

d) Seizures in advanced disease

e) CT normal for age

IV. Features that make the diagnosis of PROBABLE Alzheimer's disease uncertain or unlikely:

a) Sudden, apoplectic onset

b) Focal neurologic findings such as hemiparesis, sensory loss, visual field deficits, and incoordination early in the course of the illness

c) Seizures or gait disturbances at the onset or very early in the course of the illness

Probable Alzheimer´s disease if:

- all criteria I are answered with "yes"

POSSIBLE ALZHEIMER´S DISEASE

- may be made on the basis of the dementia syndrome, in the absence of other neurologic, psychiatric, or systemic disorders sufficient to cause dementia, in the presence of variations in the onset, in the presentation, or in the clinical course
- may be made in the presence of a second systemic or brain disorder sufficient to produce dementia, which is not considered to be the cause of the dementia
- should be used in research studies when a single, gradually progressive severe cognitive deficit is identified in the absence of other identifiable cause

Possible Alzheimer´s disease if:

- criteria above are appropriate

DEFINITE ALZHEIMER'S DISEASE

Definite AD refers to histopathologically confirmed disease. The methods by which pathologists should make this diagnosis are described in Mirra *et al.* (Mirra *et al.*, 1991; Mirra *et al.*, 1993). Histopathological methods (and AD criteria) are imperfect and continue to be refined but in essence depend on demonstrating sufficient age-related densities of neurofibrillary tangles and amyloid plaques (Newell *et al.*, 1999).

FRONTOTEMPORAL LOBAR DEGENERATION

Criteria given by Neary *et al.* (Neary *et al.*, 1998)

I. Core diagnostic features

a) Insidious onset and gradual progression

b) Early decline in social interpersonal conduct

c) Early impairment in regulation of personal conduct

d) Early emotional blunting

e) Early loss of insight

II. Supportive diagnostic features

a) Behavioural disorder

(i) Decline in personal hygiene and grooming

(ii) Mental rigidity and inflexibility

(iii) Distractibility and impersistence

(iv) Hyperorality and dietary changes

(v) Perseverative and stereotyped behaviour

(vi) Utilization behaviour

b) Speech and Language

(i) Altered speech output

Aspontaneity and economy of speech

Press of speech

(ii) Stereotypy of speech

(iii) Echolalia

(iv) Perseveration

(v) Mutism

c) Physical signs

(i) Early primitive reflexes

(ii) Early incontinence

(iii) Akinesia, rigidity, and tremor

(iv) Low and labile blood pressure

d) Investigations

(i) Neuropsychology: significant impairment on frontal lobe tests in the absence of severe amnesia, aphasia, or perceptuospatial disorder

(ii) Electroencephalography: normal on conventional EEG despite clinically evident dementia

(iii) Brain imaging (structural and/or functional): predominant frontal and/or anterior temporal abnormality

Extension list for Frontotemporal Dementia, Progressive Nonfluent Aphasia, and Semantic Dementia

III. Supportive Features

a) Onset before 65: positive family history of similar disorder in first-degree relative

b) Bulbar palsy, muscular weakness and wasting, fasciculations (associated motor neuron disease present in a minority of patients)

IV. Diagnostic exclusion features

a) Historical and clinical

(i) Aprupt onset with ictal events

(ii) Head trauma related to onset

(iii) Early, severe amnesia

(iv) Spatial disorientation

(v) Festinant speech with loss of train of thought

(vi) Myoclonus

(vii) Corticospinal weakness

(viii) Cerebellar ataxia

(ix) Choreoathetosis

b) Investigation

(i) Brain imaging: predominant postcentral structural or functional deficit, multifocal lesions on CT or MRI

(ii) Laboratory tests indicating brain involvement or inflammatory disorder such as MS, syphilis, AIDS and herpes simplex encephalitis.

V. Relative diagnostic exclusion features

a) Typical history of chronic alcoholism

b) Sustained hypertension

c) History of vascular disease (such as angina, claudication).

FTLD if:

- all core diagnostic features for FTLD (I a-e) are answered with "yes"

AND

none of the exclusion criteria under IV is answered with "yes"

DEMENTIA WITH LEWY BODIES

Criteria given by McKeith *et al.* (McKeith *et al.*, 1996)

1. Central feature required for a diagnosis of DLB

a) Progressive cognitive decline

b) Cognitive decline interferes with normal social or occupational function.

2. Core features

a) Fluctuating cognition with pronounced variations in attention and alertness

b) Recurrent visual hallucinations that are typically well formed and detailed

c) Spontaneous motor features of parkinsonism

3. A diagnosis of DLB is less likely in the presence of

a) Stroke disease, evident as focal neurologic signs or on brain imaging

b) Evidence on physical examination and investigation of any physical illness or other brain disorder sufficient to account for the clinical picture.

Possible Lewy Body Dementia if:

- both criteria 1 are answered with "yes"

AND

- none of criteria 3 are answered with "yes"

AND

- at least one of the core features 2 are answered with "yes"

Probable Lewy Body Dementia if:

- both criteria 1 are answered with "yes"

AND

- none of criteria 3 are answered with "yes"

AND

- two of the core features 2 are answered with "yes"

VASCULAR DEMENTIA

Criteria given by the National Institute of Neurological Disorders and Stroke - Association Internationale pour la Recherché et L'Enseignement en Neurosciences (NINDS-AIREN) (Roman *et al.*, 1993)

I. Criteria for the clinical diagnosis of PROBABLE vascular dementia include all of the following:

a) Dementia defined by cognitive decline from a previously higher level of functioning and manifested by impairment of memory and of two or more cognitive domains (orientation, attention, language, visuospatial functions, executive functions, motor control, and praxis), preferably established by clinical examination and documented by neuropsychological testing; deficits should be severe enough to interfere with activities of daily living not due to physical effects of stroke alone.

Exclusion criteria: cases with disturbance of consciousness, delirium, psychosis, severe aphasia, or major sensorimotor impairment precluding neuropsychological testing. Also excluded are systemic disorders or other brain diseases (such as AD) that in and of themselves could account for deficits in memory and cognition.

b) Cerebrovascular disease (CVD), defined by the presence of focal signs on neurologic examination, such as hemiparesis, lower facial weakness, Babinski sign, sensory deficit, hemianopia, and dysarthria consistent with stroke (with or without history of stroke), and evidence of relevant CVD by brain imaging (CT or MRI) including multiple large-vessel infarcts or a single strategically placed infarct (angular gyrus, thalamus, basal forebrain, or PCA or ACA territories), as well as multiple basal ganglia and white matter lacunes or extensive periventricular white matter lesions, or combinations thereof.

c) A relationship between the above two disorders, manifested or inferred by the presence of one or more of the following:

(i) onset of dementia within three 3 months following a recognised stroke

(ii) abrupt deterioration in cognitive functions; or fluctuating, stepwise progression of cognitive deficits.

II. Clinical features consistent with the diagnosis of PROBABLE vascular dementia include the following:

a) Early presence of a gait disturbance (small-step gait or marche à petits pas, or magnetic, apraxic-ataxic or parkinsonian gait)
b) History of unsteadiness and frequent, unprovoked falls
c) Early urinary frequency, urgency, and other urinary symptoms not explained by urologic disease
d) Pseudobulbar palsy
e) Personality and mood changes, abulia, depression, emotional incontinence, or other subcortical deficits including psychomotor retardation and abnormal executive function.

III. Features that make the diagnosis of vascular dementia uncertain or unlikely include
a) Early onset of memory deficit and progressive worsening of memory and other cognitive functions such as language (transcortical sensory aphasia), motor skills (apraxia), and perception (agnosia), in the absence of corresponding focal lesions on brain imaging;
b) Absence of focal neurologic signs, other than cognitive disturbance; and
c) Absence of cerebrovascular lesions on brain CT or MRI.

<u>MCI</u>
Criteria given by Petersen *et al.* (Petersen *et al.*, 1999)
a) Memory complaint
b) Normal activities of daily living
c) Normal general cognitive function
d) Abnormal memory for age, and
e) Not demented
f) Absence of psychiatric symptoms

Appendix Two: MINI MENTAL STATE EXAMINATION

Mini Mental State Examination (MMSE) described by Folstein *et al.* (Folstein *et al.*, 1975)

Appendix Three: IMAGE ACQUISITION PROTOCOLS

All volumetric scans were acquired on a 1.5 Tesla Signa unit (General Electric, Waukesha, WI, USA) as 124 contiguous 1.5mm coronal slices using one of the two following acquisitions:

Standard imaging protocol

Scans were performed by the staff at St Mary's Hospital, Paddington and the Queen Square Imaging Centre. These were taken using a 256x128 image matrix with the field of view being 24cm (acquisition parameters: repetition time=35ms; echo time=5ms; flip angle=35°).

MIRIAD protocol

Scans were acquired by Dave MacManus of the NMR group, Queen Square. This protocol was an inversion recovery prepared fast SPGR sequence with a 256x256 image matrix and the field of view being 24cm (acquisition parameters: repetition time=15ms; echo time=5.4ms; flip angle=15°; inversion time=650ms).

Appendix Four: VOLUMETRIC ANALYSIS

Volumetric analysis was performed using the MIDAS software, following the protocols detailed below. All segmentations were performed on T1-weighted images. Reproducibility error was estimated as the mean absolute percentage difference between repeat segmentations on the same scan and/or an ICC calculated from the same measurements.

Whole brain

Whole brain segmentation was performed using a protocol described in detail by Freeborough *et al.* (Freeborough *et al.*, 1997). This is a semi-automated algorithm which identifies voxels within the brain using interactive thresholding. Brain tissue was then isolated from the surrounding tissue such as scalp and dura using a series of erosions and dilations (see Appendix Figure 1). All segmentations were performed by Rachael Scahill, Rhian Jenkins, Jennifer Whitwell, Shona Price, Valerie Anderson and myself. Inter-rater reproducibility was 0.42% and intra-rater reproducibility was 0.36%.

Total Intracranial Volume

Total Intracranial Volume (TIV) was delineated using a protocol described by Whitwell *et al.* (Whitwell *et al.*, 2001). A lower threshold of 30% of the mean intracranial signal intensity was set to outline the outer border of dura, with manual editing where necessary. The inferior boundary was set as the lowest slice in which cerebellar tissue was present. An area measurement was performed on every tenth axial slice, and a volume estimate was obtained using linear interpolation. This is illustrated in Appendix Figure 2. All volumetric measurements were performed by Jennifer Whitwell, and intra-rater reproducibility was 0.23%.

Hippocampus

See Appendix Figure 3. The hippocampus was delineated by reference to the neuroanatomical atlas by Duvernoy (Duvernoy, 1998). The caudal most measurement of the hippocampus was taken at the level of the crus of the fornix. The rostral boundary of the hippocampus was at its junction with the amygdala. Superiorly, medially and laterally the hippocampus was bounded by the ambient cistern and temporal horn of the lateral ventricle, and inferiorly by the subjacent white matter of the subiculum. This measure included the dentate gyrus, the hippocampus proper and the alveus. The

hippocampal tail was excluded in order to achieve satisfactory reproducibility of segmentation, as recommended by Watson *et al.* (Watson *et al.*, 1992). A relative-to-brain threshold of 70% was used to aid delineation of grey matter from CSF. All segmentations were performed by Rachael Scahill, Basil Ridha, Valerie Anderson and myself. Inter-rater reproducibility error was 5% and intra-rater reproducibility error was 3% (an ICC of 0.98).

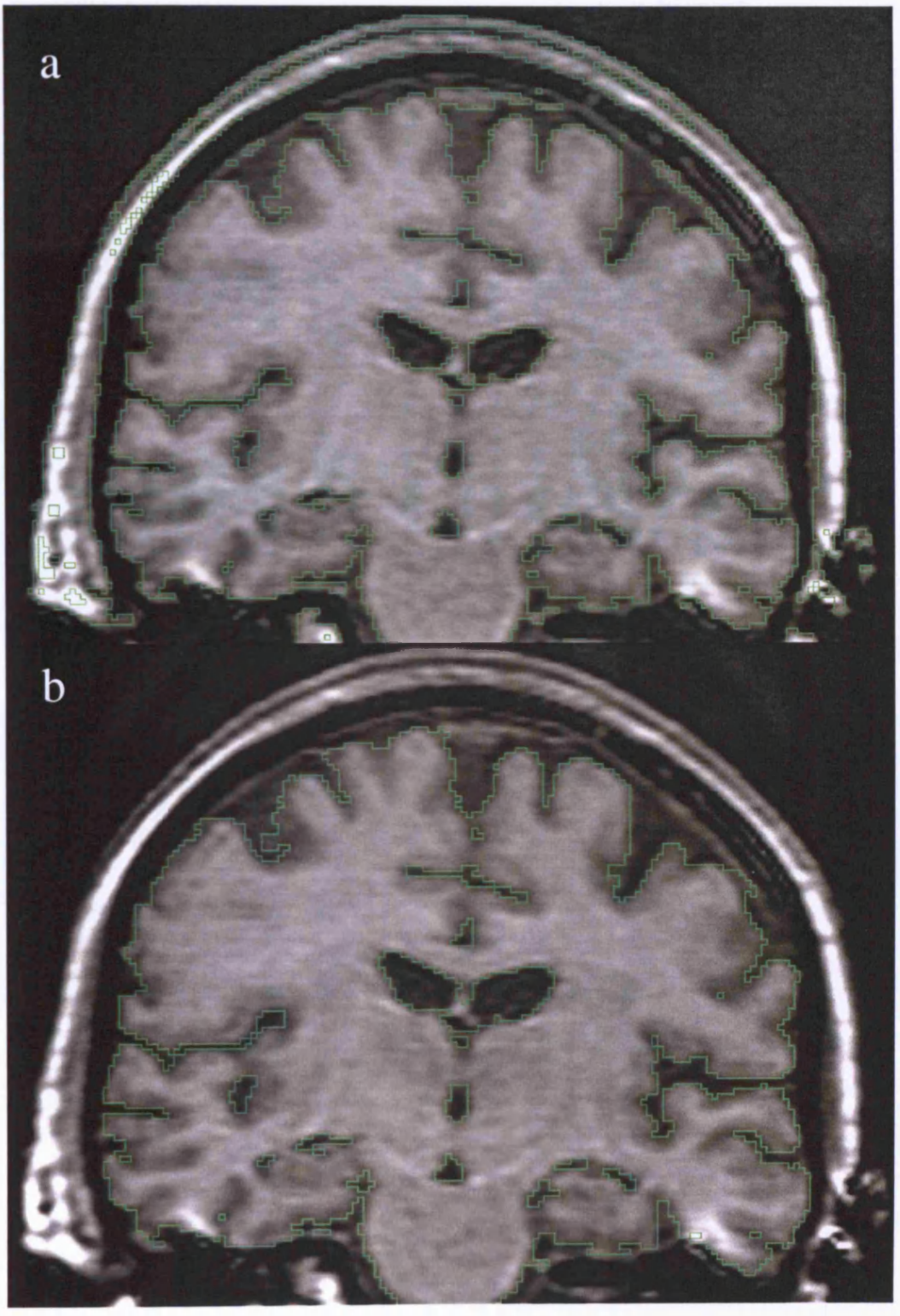
Cingulate gyrus

See Appendix Figure 4. Cingulate segmentation was performed using the protocol (Jones *et al,* 2006). The posterior limit of the cingulate gyrus was taken as the splenial sulcus, the superior and anterior limit as the bottom of the cingulate sulcus and the inferior limit as the corpus collosum. The cingulate was outlined in the sagittal view (a), from medial to lateral parts using strict relative-to-brain thresholds of 70% and 110% to aid delineation of grey matter from CSF and white matter respectively. The cingulate region was then edited coronally (b) and (c), with (b) showing the coronal slice at the posterior dashed line in (a), and (c) showing the more anterior coronal slice shown by the anterior dashed line in (a). Subdivisions were manually determined using the anterior and posterior comissures as cut-offs (arrows on (d) showing mid-sagittal "slice"). Cingulate gyrus anterior to the anterior commissure was considered to be rostral anterior cingulate (RAC), cingulate posterior to the posterior commissure is posterior cingulate (PC) and cingulate between these regions was caudal anterior cingulate (CAC). Segmentations were performed by Bethany Jones and myself, and the intra-rater reproducibility error of whole cingulate was 4% (a reliability coefficient of 0.98), and the inter-rater reproducibility error of whole cingulate was 5% (an ICC of 0.98).

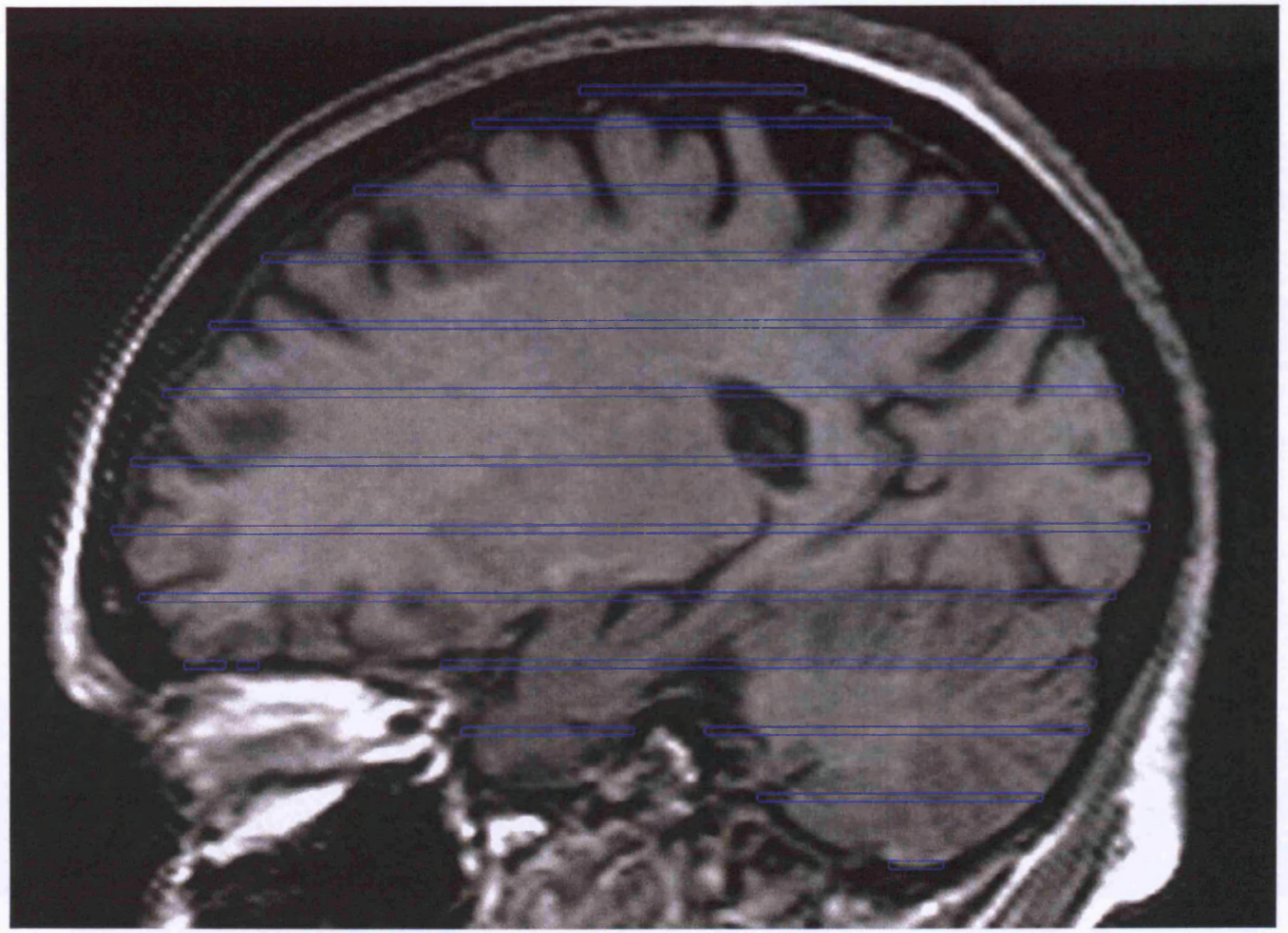
Amygdala

See Appendix Figure 5. Amygdala segmentation was performed using a modification of a previously published protocol (Sheline *et al.*, 1998). The rostral limit corresponds to the rostral extreme of the temporal stem. Caudally the amygdala was defined using the border with the lateral ventricle and the alveus, ending when the alveus disappears. Superiorly the boundary was set anteriorly by a line connecting the inferior point of the lateral fissure to the lateral most point of the paramedian cisternae, and posteriorly by the superior and lateral borders of the optic tract. Infero-laterally the amygdala was bounded by white matter and medially by the ambient cistern. All amygdala segmentations were performed by Jenny Whitwell, and the intra-rater reproducibility error was 4% (an ICC of 0.98).

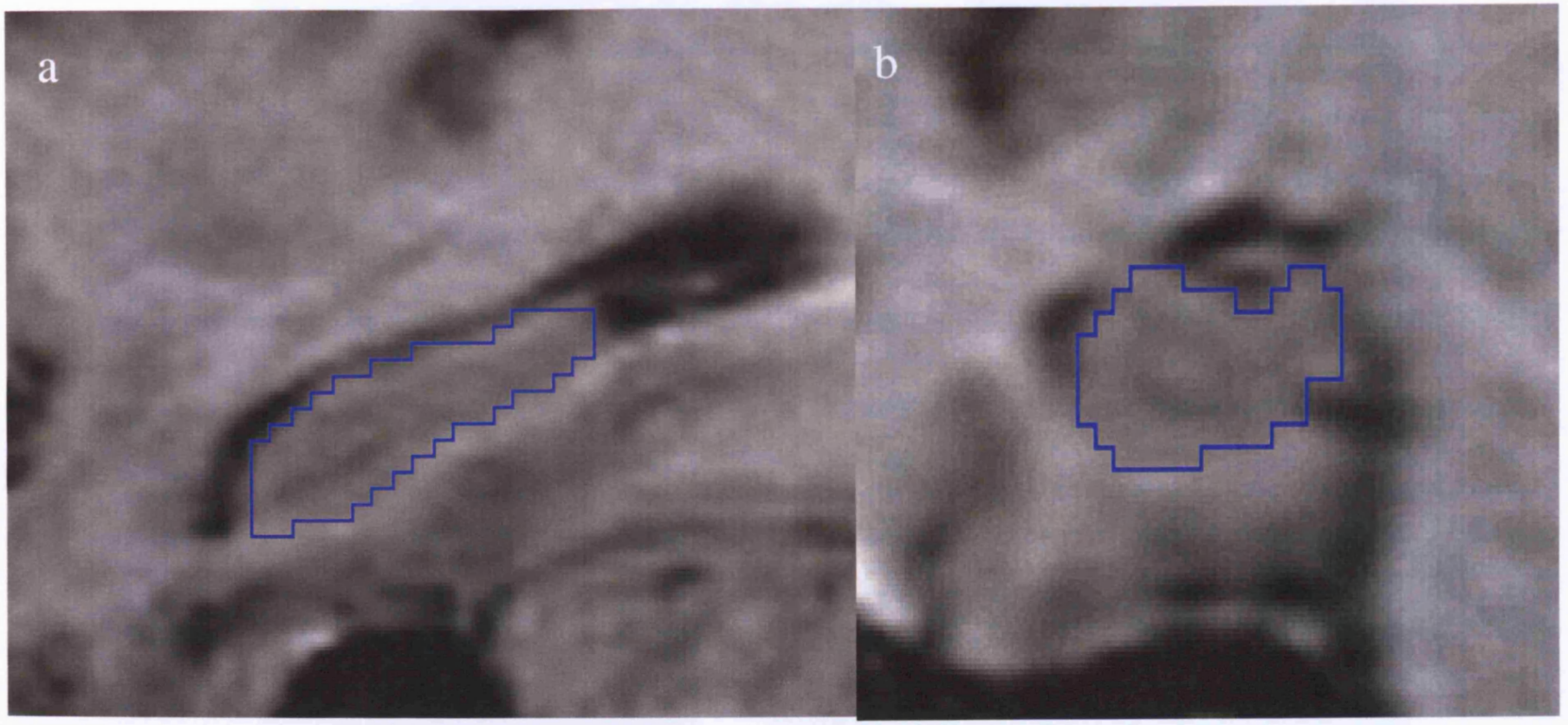
Appendix Figure 1: Whole brain segmentation in coronal view.
a) shows intensity thresholding and b) brain delineation following erosions and dilations



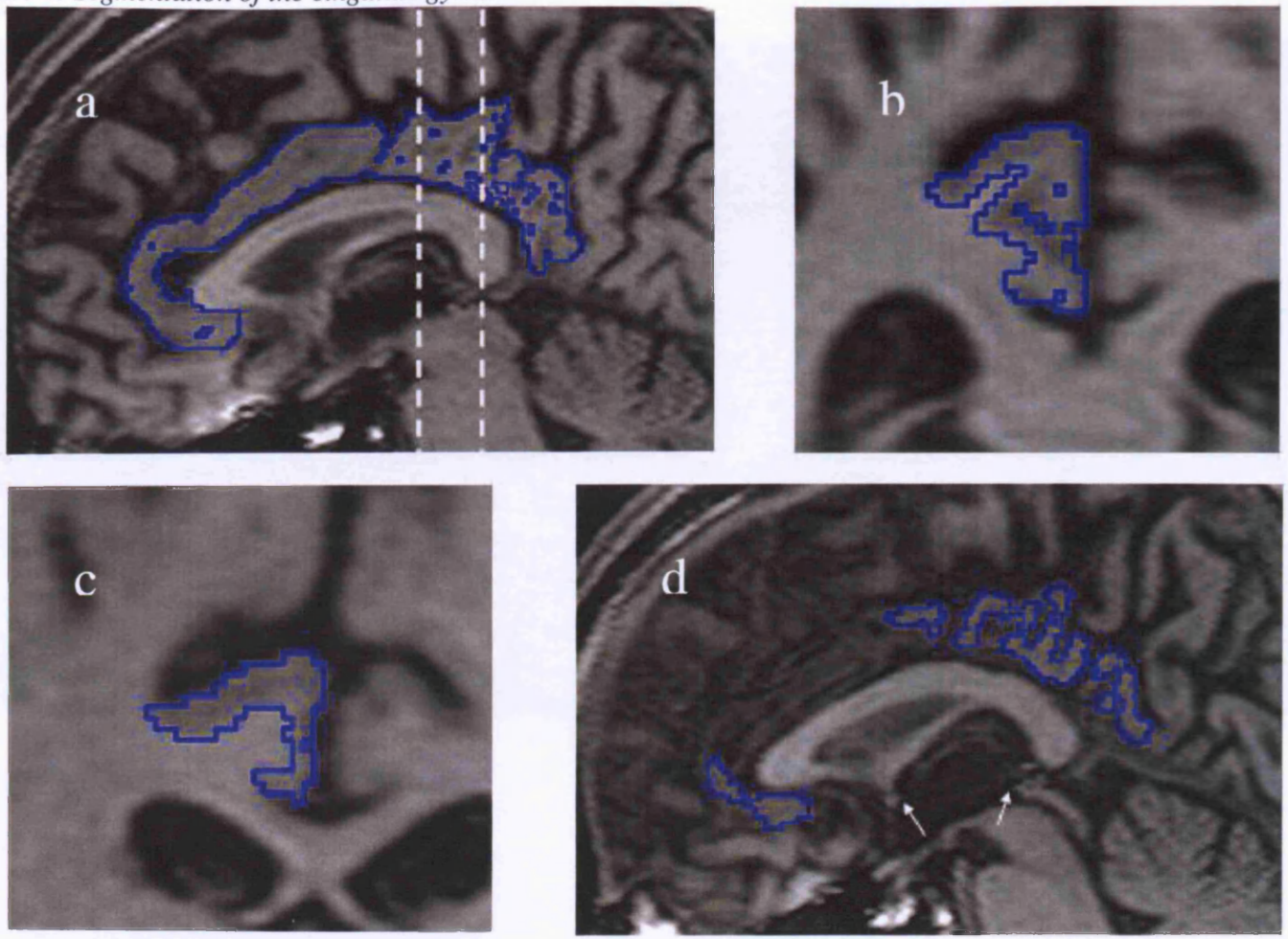
Appendix Figure 2: TIV segmentation in sagittal view.



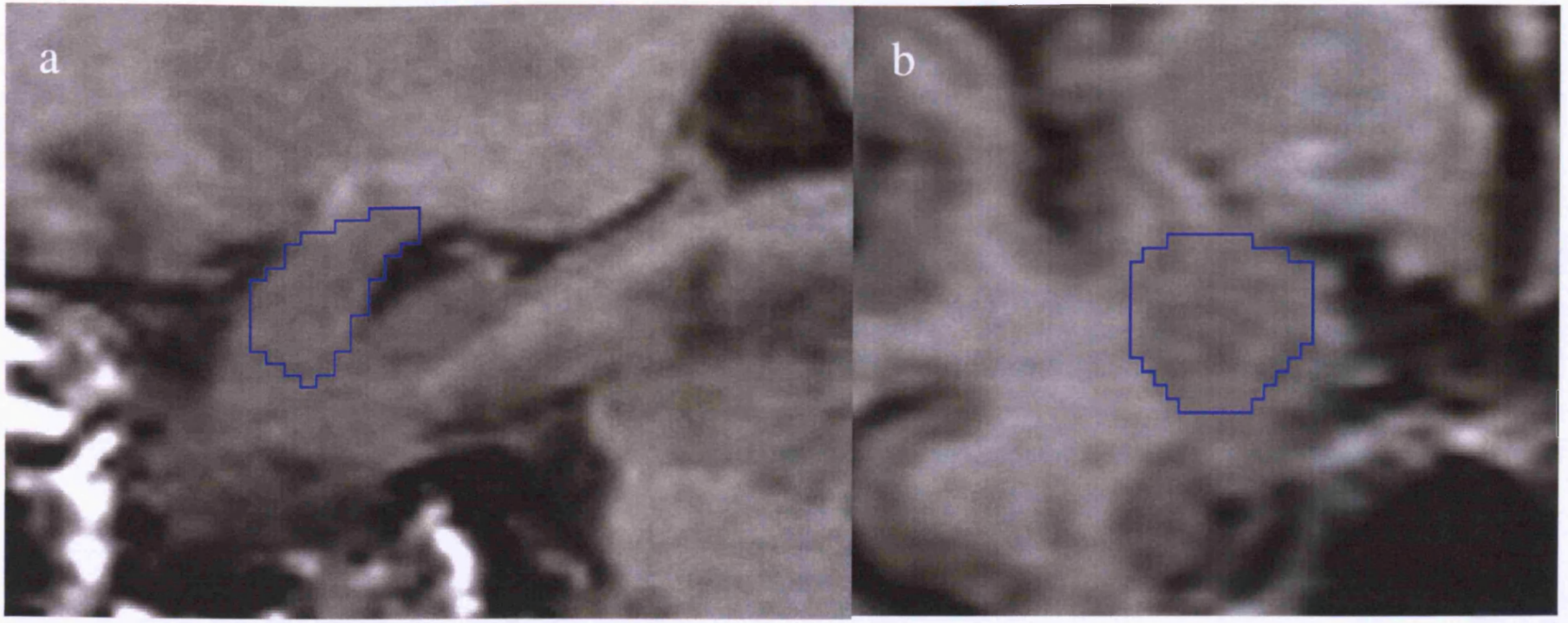
Appendix Figure 3: Hippocampal segmentation.



Appendix Figure 4: Segmentation of the cingulate gyrus.



Appendix Figure 5: Segmentation of the amygdala.



GLOSSARY

AC: Anterior cingulate

AD: Alzheimer's disease

APOE: Apolipoprotein gene

APP: Amyloid precursor protein

BBSI: Brain boundary shift integral

BSI: Boundary shift integral

CAC: Caudal anterior cingulate

CSF: Cerebrospinal fluid

CT: Computed tomography

DLB: Dementia with Lewy bodies

dof: degrees of freedom

EEG: Electroencephalogram

FAD: Familial Alzheimer's disease

fMRI: Functional Magnetic Resonance Imaging

FTLD: Fronto-temporal lobar degeneration, referring to the spectrum of diseases including fronto-temporal dementia, semantic dementia and primary progressive aphasia. The terminology covering these conditions is not always clear, and some authors use the term FTD to refer to the whole group, and others only to the frontal variant of this disorder.

fvFTD: Fronto-temporal dementia. This refers to the frontal variant of FTLD.

MCI: Mild cognitive impairment

MIDAS: Medical Information Display and Analysis System

MIRIAD: Minimal Interval Resonance Imaging in AD

MMSE: Mini Mental State Examination

MRI: Magnetic Resonance Imaging

MRS: Magnetic Resonance Spectroscopy

MTL: Medial temporal lobe

NFT: Neurofibrillary tangle

PC: Posterior cingulate

PET: Positron Emission Tomography

PNFA: Progressive non-fluent aphasia, a clinical sub-type of FTLD

RAC: Rostral anterior cingulate

ROI: Region of interest

RTLA: Right temporal lobe atrophy, a sub-type of FTLD

SAD: Sporadic Alzheimer's disease

SD: Semantic dementia, a clinical sub-type of FTLD, and standard deviations in a statistical context.

SPECT: Single photon emission computed tomography

SPM: Statistical parametric mapping, the software package used for statistical analysis of structural and functional images

TE: Time to echo i.e. time between transmission of RF pulse and collection of signal

TIV: Total intracranial volume – a measure of total capacity within the skull, used as a maximum estimate of pre-morbid brain volume

TR: Time to repeat i.e. time between RF pulses

VBM: Voxel-based morphometry, the application of SPM to highlight grey matter differences between groups, based on structural MRIs.

BIBLIOGRAPHY

1. Alexander, G. E., Chen, K., Pietrini, P., Rapoport, S.I., Reiman, E.M., 2002. Longitudinal PET Evaluation of Cerebral Metabolic Decline in Dementia: A Potential Outcome Measure in Alzheimer's Disease Treatment Studies. American Journal of Psychiatry 159, 738-745.

2. American Psychiatric Association. 1994. *Diagnostic and Statistical Manual of Mental Disorders. (4th edn) (DSM-IV).* American Psychiatric Association. Washington DC.

3. Andreasen, N. C., Rajarethinam, R., Cizadlo, T., Arndt, S., Swayze, V.W., Flashman, L.A., O'Leary, D.S., Ehrhardt, J.C., Yuh, W.T.C., 1996. Automatic Atlas-Based Volume Estimation of Human Brain Regions from MR Images. J Comput Assist Tomogr 20, 98-106.

4. Arnaiz, E., Almkvist, O., 2003. Neuropsychological features of mild cognitive impairment and preclinical Alzheimer's disease. Acta Neurol Scand Suppl 179, 34-41.

5. Arnold, S. E., Hyman, B.T., Flory, J., Damasio, A.R., Van Hoesen, G.W., 1991. The topographical and neuroanatomical distribution of neurofibrillary tangles and neuritic plaques in the cerebral cortex of patients with Alzheimer's disease. Cereb.Cortex 1, 103-116.

6. Ashburner, J., Csernansky, J.G., Davatzikos, C., Fox, N.C., Frisoni, G.B., Thompson, P.M., 2003. Computer-assisted imaging to assess brain structure in healthy and diseased brains. Lancet Neurology 2, 79-88.

7. Ashburner, J., Friston, K.J., 1999. Nonlinear spatial normalization using basis functions. Hum.Brain Mapp. 7, 254-266.

8. Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry--the methods. Neuroimage 11, 805-821.

9. Ashtari, M., Greenwald, B.S., Kramer-Ginsberg, E., Hu, J., Wu, H., Patel, M., Aupperle, P., Pollack, S., 1999. Hippocampal amygdala volumes in geriatric depression. Psychol Med 29, 629-638.

10. Ashton, E. A., Parker, K.J., Berg, M.J., Chen, C.W., 1997. A novel volumetric feature extraction technique with applications to MR images. IEEE Trans.Med Imaging 16, 365-371.

11. Baleydier, C., Mauguiere, F., 1980. The duality of the cingulate gyrus in monkey. Neuroanatomical study and functional hypothesis. Brain 103, 525-554.

12. Ballmaier, M., O'Brien, J.T., Burton, E.J., Thompson, P.M., Rex, D.E., Narr, K.L., McKeith, I.G., Deluca, H., Toga, A.W., 2004. Comparing gray matter loss profiles between dementia with Lewy bodies and Alzheimer's disease using cortical pattern matching: diagnosis and gender effects. Neuroimage 23, 325-335.

13. Barber, R., Ballard, C., McKeith, I.G., Gholkar, A., O'Brien, J.T., 2000. MRI volumetric study of dementia with Lewy bodies: A comparison with AD and vascular dementia. Neurology 54, 1304-1309.

14. Barnes, J., Scahill, R.I., Schott, J.M., Frost, C., Rossor, M.N., Fox, N.C., 2005. Does Alzheimer's disease affect hippocampal asymmetry? Evidence from a cross-sectional and longitudinal volumetric MRI study. Dement Geriatr Cogn Disord 19, 338-344.

15. Baron, J. C., Chetelat, G., Desgranges, B., Perchey, G., Landeau, B., De La Sayette, V., Eustache, F., 2001. *In vivo* mapping of gray matter loss with voxel-based morphometry in mild Alzheimer's disease. Neuroimage 14, 298-309.

16. Bartko, J. J., 1966. The intraclass correlation coefficient as a measure of reliability. Psychol Rep. 19, 3-11.

17. Benveniste, H., Einstein, G., Kim, K.R., Hulette, C., Johnson, G.A., 1999. Detection of neuritic plaques in Alzheimer's disease by magnetic resonance

microscopy. Proceedings of the National Academy of Sciences of the United States of America 96, 14079-14084.

18. Bhatia, S., Bookheimer, S.Y., Gaillard, W.D., Theodore, W.H., 1993. Measurement of whole temporal lobe and hippocampus for MR volumetry: Normative data. Neurology 43, 2006-2010.

19. Bigler, E. D., Tate, D.F., Miller, M.J., Rice, S.A., Hessel, C.D., Earl, H.D., Tschanz, J.T., Plassman, B., Welsh-Bohmer, K.A., 2002. Dementia, asymmetry of temporal lobe structures, and apolipoprotein E genotype: relationships to cerebral atrophy and neuropsychological impairment. J Int Neuropsychol Soc 8, 925-933.

20. Bobinski, M., de Leon, M.J., Wegiel, J., DeSanti, S., Convit, A., Saint Louis, L.A., Rusinek, H., Wisniewski, H.M., 2000. The histological validation of post mortem magnetic resonance imaging-determined hippocampal volume in Alzheimer's disease. Neuroscience 95, 721-725.

21. Boccardi, M., Pennanen, C., Laakso, M., Testa, C., Geroldi, C., Soininen, H., Frisoni, G.B., 2002. Amygdaloid atrophy in frontotemporal dementia and Alzheimer's disease. Neurosci.Lett. 335, 139-143.

22. Boccardi, M., Sabattoli, F., Laakso, M.P., Testa, C., Rossi, R., Beltramello, A., Soininen, H., Frisoni, G.B., 2005. Frontotemporal dementia as a neural system disease. Neurobiol Aging 26, 37-44.

23. Boccardi, M., Laakso, M.P., Bresciani, L., Galluzzi, S., Geroldi, C., Beltramello, A., Soininen, H., Frisoni, G.B., 2003. The MRI pattern of frontal and temporal brain atrophy in fronto-temporal dementia. Neurobiol Aging 24, 95-103.

24. Boxer, A. L., Rankin, K.P., Miller, B.L., Schuff, N., Weiner, M., Gorno-Tempini, M.L., Rosen, H.J., 2003. Cinguloparietal atrophy distinguishes Alzheimer disease from semantic dementia. Arch.Neurol 60, 949-956.

25. Boyes, R.G., Barnes, J, Fox. N.C., 2004. Partitioning the cingulate using Bezier curves. MIUA, 208-211.

26. Braak, E. and Braak, H. 1993. "Alzheimer neuropathology and limbic circuits." In B.A.Vogt and M.Gabriel, editors, *Neurobiology of cingulate cortex and limbic thalamus: a comprehensive handbook*. Birkhauser. Boston. 606-627.

27. Braak, H., Braak, E., 1991. Neuropathological staging of Alzheimer-related changes. Acta Neuropathologica 82, 239-259.

28. Braak, H., Braak, E., Bohl, J., 1993. Staging of Alzheimer-related cortical destruction. European Neurology 33, 403-408.

29. Bradley, K. M., Bydder, G.M., Budge, M.M., Hajnal, J.V., White, S.J., Ripley, B.D., Smith, A.D., 2002. Serial brain MRI at 3-6 month intervals as a surrogate marker for Alzheimer's disease. Br.J.Radiol. 75, 506-513.

30. Bussiere, T., Friend, P.D., Sadeghi, N., Wicinski, B., Lin, G.I., Bouras, C., Giannakopoulos, P., Robakis, N.K., Morrison, J.H., Perl, D.P., Hof, P.R., 2002. Stereologic assessment of the total cortical volume occupied by amyloid deposits and its relationship with cognitive status in aging and Alzheimer's disease. Neuroscience 112, 75-91.

31. Callen, D. J. A., Black, S.E., Gao, F., Caldwell, C.B., Szalai, J.P., 2001. Beyond the hippocampus. MRI volumetry confirms widespread limbic atrophy in AD. Neurology 57, 1669-1674.

32. Cardenas, V. A., Du, A.T., Hardin, D., Ezekiel, F., Weber, P., Jagust, W.J., Chui, H.C., Schuff, N., Weiner, M.W., 2003. Comparison of methods for measuring longitudinal brain change in cognitive impairment and dementia. Neurobiol Aging 24, 537-544.

33. Carmichael, O. T., Aizenstein, H.A., Davis, S.W., Becker, J.T., Thompson, P.M., Meltzer, C.C., Liu, Y., 2005. Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. Neuroimage 27, 979-990.

34. Chan, D., Fox, N., Rossor, M., 2002. Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. Neurology 58, 838.

35. Chan, D., Fox, N.C., Jenkins, R., Scahill, R.I., Crum, W.R., Rossor, M.N., 2001a. Rates of global and regional cerebral atrophy in AD and frontotemporal dementia. Neurology 57, 1756-1763.

36. Chan, D., Fox, N.C., Scahill, R.I., Crum, W.R., Whitwell, J.L., Leschziner, G., Rossor, A.M., Stevens, J.M., Cipolotti, L., Rossor, M.N., 2001b. Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. Ann.Neurol. 49, 433-442.

37. Chan, D., Janssen, J.C., Whitwell, J.L., Watt, H.C., Jenkins, R., Frost, C., Rossor, M.N., Fox, N.C., 2003. Change in rates of cerebral atrophy over time in early-onset Alzheimer's disease: longitudinal MRI study. Lancet 362, 1121-1122.

38. Chetelat, G., Baron, J.C., 2003. Early diagnosis of Alzheimer's disease: contribution of structural neuroimaging. Neuroimage 18, 525-541.

39. Christensen, G. E., Rabbitt, R.D., Miller, M.I., 1996. Deformable templates using large deformation kinematics. IEEE Trans Image Proc 5, 1435-1447.

40. Collins, D. L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J Comput.Assist.Tomogr. 18, 192-205.

41. Convit, A., de Leon, M.J., Golomb, J., George, A.E., Tarshish, C.Y., Bobinski, M., Tsui, W., De Santi, S., Wegiel, J., Wisniewski, H., 1993. Hippocampal atrophy in early Alzheimer's disease: anatomic specificity and validation. Psychiatric Quarterly 64, 371-387.

42. Cook, M. J., Fish, D.R., Shorvon, S.D., Straughan, K., Stevens, J.M., 1992. Hippocampal volumetric and morphometric studies in frontal and temporal lobe epilepsy. Brain 115, 1001-1015.

43. Crum, W. R., Griffin, L.D., Hill, D.L., Hawkes, D.J., 2003. Zen and the art of medical image registration: correspondence, homology, and quality. Neuroimage 20, 1425-1437.

44. Crum, W. R., Hartkens, T., Hill, D.L.G., 2004a. Non-rigid image registration: theory and practice. Br J Radiol 77, 140-153.

45. Crum, W. R., Rueckert, D., Jenkinson, M., Kennedy, D., Smith, S.M., 2004b. A framework for detailed objective comparison of non-rigid registration algorithms in neuroimaging. Medical Image Computing and Computer-Assisted Intervention - Miccai 2004, Pt 1, Proceedings 3216, 679-686.

46. Crum, W. R., Scahill, R.I., Fox, N.C., 2001. Automated hippocampal segmentation by regional fluid registration of serial MRI: validation and application in Alzheimer's disease. Neuroimage 13, 847-855.

47. Csernansky, J. G., Joshi, S., Wang, L., Haller, J.W., Gado, M., Miller, J.P., Grenander, U., Miller, M.I., 1998. Hippocampal morphometry in schizophrenia by high dimensional brain mapping. Proceedings of the National Academy of Sciences of the United States of America 95, 11406-11411.

48. Csernansky, J. G., Wang, L., Joshi, S., Miller, J.P., Gado, M., Kido, D., McKeel, D., Morris, J.C., Miller, M.I., 2000. Early DAT is distinguished from aging by high-dimensional mapping of the hippocampus. Neurology 55, 1636-1643.

49. Csernansky, J. G., Wang, L., Swank, J., Miller, J.P., Gado, M., McKeel, D., Miller, M.I., Morris, J.C., 2005. Preclinical detection of Alzheimer's disease: hippocampal shape and volume predict dementia onset in the elderly. Neuroimage 25, 783-792.

50. Cuenod, C. A., Denys, A., Michot, J.L., Jehenson, P., Forette, F., Kaplan, D., Syrota, A., Boller, F., 1993. Amygdala atrophy in Alzheimer's disease. An in vivo magnetic resonance imaging study. Archives of Neurology 50, 941-945.

51. Dawbarn, D. and Allen, S.J., 2001. *Neurobiology of Alzheimer's Disease.* Oxford University Press. Oxford.

52. de Leon, M.J., George, A.E., Stylopoulos, L.A., Miller, D.C., 1994. "Computed tomographic hippocampal atrophy: an early marker for Alzheimer's disease." *Progress in Dementia research.* 55-62.

53. de Leon, M. J., George, A.E., Stylopoulos, L.A., Smith, G., Miller, D.C., 1989. Early marker for Alzheimer's disease: The atrophic hippocampus. Lancet 2, 672-673.

54. DeCarli, C., Murphy, D.G., Gillette, J.A., Haxby, J.V., Teichberg, D., Schapiro, M.B., Horwitz, B., 1994. Lack of age-related differences in temporal lobe volume of very healthy adults. Am J Neuroradiol 15, 689-696.

55. Delacourte, A., David, J.P., Sergeant, N., Buee, L., Wattez, A., Vermersch, P., Ghozali, F., FalletBianco, C., Pasquier, F., Lebert, F., Petit, H., DiMenza, C., 1999. The biochemical pathway of neurofibrillary degeneration in aging and Alzheimer's disease. Neurology 52, 1158-1165.

56. Donner, A., Eliasziw, M., 1987. Sample size requirements for reliability studies. Stat.Med 6, 441-448.

57. Du, A. T., Schuff, N., Amend, D., Laakso, M.P., Hsu, Y.Y., Jagust, W.J., Yaffe, K., Kramer, J.H., Reed, B., Norman, D., Chui, H.C., Weiner, M.W., 2001. Magnetic resonance imaging of the entorhinal cortex and hippocampus in mild cognitive impairment and Alzheimer's disease. J Neurol Neurosurg Psychiat 71, 441-447.

58. Du, A. T., Schuff, N., Kramer, J.H., Ganzer, S., Zhu, X.P., Jagust, W.J., Miller, B.L., Reed, B.R., Mungas, D., Yaffe, K., Chui, H.C., Weiner, M.W., 2004. Higher atrophy rate of entorhinal cortex than hippocampus in AD. Neurology 62, 422-427.

59. Du, A. T., Schuff, N., Laakso, M.P., Zhu, X.P., Jagust, W.J., Yaffe, K., Kramer, J.H., Miller, B.L., Reed, B.R., Norman, D., Chui, H.C., Weiner, M.W., 2002. Effects of subcortical ischemic vascular dementia and AD on entorhinal cortex and hippocampus. Neurology 58, 1635-1641.

60. Du, A. T., Schuff, N., Zhu, X.P., Jagust, W.J., Miller, B.L., Reed, B.R., Kramer, J.H., Mungas, D., Yaffe, K., Chui, H.C., Weiner, M.W., 2003. Atrophy rates of entorhinal cortex in AD and normal aging. Neurology 60, 481-486.

61. Duchesne, S., Pruessner, J.C., Collins, D.L., 2002. Appearance-Based Segmentation of Medial Temporal Lobe Structures. Neuroimage 17, 515-531.

62. Duvernoy, H.M. 1998. *The Human Hippocampus*. Springer-Verlag. Heidelberg.

63. Ebel, R., 1951. Estimation of the reliability of ratings. Psychometrika 16, 407-423.

64. Evans, J. J., Heggs, A.J., Antoun, N., Hodges, J.R., 1995. Progressive prosopagnosia associated with selective right temporal lobe atrophy. A new syndrome? Brain 118, 1-13.

65. Ezekiel, F., Chao, L., Kornak, J., Du, A.T., Cardenas, V., Truran, D., Jagust, W., Chui, H., Miller, B., Yaffe, K., Schuff, N., Weiner, M., 2004. Comparisons between global and focal brain atrophy rates in normal aging and Alzheimer disease: Boundary Shift Integral versus tracing of the entorhinal cortex and hippocampus. Alzheimer Dis Assoc Disord 18, 196-201.

66. Ferrer, I., Boada, R.M., Sanchez Guerra, M.L., Rey, M.J., Costa-Jussa, F., 2004. Neuropathology and pathogenesis of encephalitis following amyloid-beta immunization in Alzheimer's disease. Brain Pathol. 14, 11-20.

67. Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der, K.A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341-355.

68. Fleis, J.L. 1986. *Design and analysis of clinical experiments*. Wiley-Interscience.

69. Folstein, M., Folstein, S., McHughs, P., 1975. The "Mini mental state": a practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research 12, 189-198.

70. Fox, N. C., 1999. Magnetic Resonance Imaging in Alzheimer's Disease: from diagnosis to measuring therapeutic effect. Alzheimer's Report 2, 5-12.

71. Fox, N. C., Black, R.S., Gilman, S., Rossor, M.N., Griffith, S.G., Jenkins, L., Koller, M., 2005. Effects of A beta immunization (AN1792) on MRI measures of cerebral volume in Alzheimer disease. Neurology 64, 1563-1572.

72. Fox, N. C., Cousens, S., Scahill, R., Harvey, R.J., Rossor, M.N., 2000. Using serial registered brain magnetic resonance imaging to measure disease progression in Alzheimer disease: power calculations and estimates of sample size to detect treatment effects. Arch.Neurol. 57, 339-344.

73. Fox, N. C., Crum, W.R., Scahill, R.I., Stevens, J.M., Janssen, J.C., Rossor, M.N., 2001. Imaging of onset and progression of Alzheimer's disease with voxel-compression mapping of serial magnetic resonance images. Lancet 358, 201-205.

74. Fox, N. C., Freeborough, P.A., 1997. Brain atrophy progression measured from registered serial MRI: validation and application to Alzheimer's disease. J Magn.Reson.Imaging 7, 1069-1075.

75. Fox, N. C., Freeborough, P.A., Rossor, M.N., 1996a. Visualisation and quantification of atrophy in Alzheimer's disease. Lancet 348, 94-97.

76. Fox, N. C., Schott, J.M., 2004. Imaging cerebral atrophy: normal ageing to Alzheimer's disease. Lancet 363, 392-394.

77. Fox, N. C., Warrington, E.K., Freeborough, P.A., Hartikainen, P., Kennedy, A.M., Stevens, J.M., Rossor, M.N., 1996b. Presymptomatic hippocampal atrophy in Alzheimer's disease: a longitudinal MRI study. Brain 119, 2001-2007.

78. Fox, N. C., Warrington, E.K., Seiffer, A.S., Agnew, S.K., Rossor, M.N., 1998. Presymptomatic cognitive deficits in individuals at risk of familial Alzheimer's disease. A longitudinal prospective study. Brain 121, 1631-1639.

79. Free, S. L., Bergin, P.S., Fish, D.R., Cook, M.J., Shorvon, S.D., Stevens, J.M., 1995. Methods for normalization of hippocampal volumes measured with MR. American Journal of Neuroradiology 16, 637-643.

80. Freeborough, P. A., Fox, N.C., 1997. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. IEEE Trans Med Imaging 16, 623-629.

81. Freeborough, P. A., Fox, N.C., 1998. Modeling brain deformations in Alzheimer's disease by fluid registration of serial 3D MR images. J Comput Assist Tomogr 22, 838-843.

82. Freeborough, P.A., Fox N.C., Kitney R.I., 1996a. "Accurate segmentation of 3D brain scans: Interactive software and algorithms." In R.Raby and D.Vicars, editors, *Proceedings of the Eurographics 14th Annual Conference (UK Chapter)*. 261-270.

83. Freeborough, P. A., Fox, N.C., Kitney, R.I., 1997. Interactive algorithms for the segmentation and quantitation of 3-D MRI brain scans. Computer Methods and Programs in Biomedicine 53, 15-25.

84. Freeborough, P. A., Woods, R.P., Fox, N.C., 1996b. Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders. J Comput Assist Tomogr 20, 1012-1022.

85. Frisoni, G. B., Testa, C., Zorzan, A., Sabattoli, F., Beltramello, A., Soininen, H., Laakso, M.P., 2002a. Detection of grey matter loss in mild Alzheimer's disease with voxel based morphometry. J Neurol Neurosurg Psychiatry 73, 657-664.

86. Frisoni, G. B., Beltramello, A., Geroldi, C., Weiss, C., Bianchetti, A., Trabucchi, M., 1996a. Brain atrophy in frontotemporal dementia. J Neurol Neurosurg Psychiatry 61, 157-165.

87. Frisoni, G. B., Beltramello, A., Weiss, C., Geroldi, C., Bianchetti, A., Trabucchi, M., 1996b. Linear measures of atrophy in mild Alzheimer disease. AJNR Am J Neuroradiol. 17, 913-923.

88. Frisoni, G. B., Beltramello, A., Weiss, C., Geroldi, C., Bianchetti, A., Trabucchi, M., 1996c. Usefulness of simple measures of temporal lobe atrophy in probable Alzheimer's disease. Dementia 7, 15-22.

89. Frisoni, G. B., Geroldi, C., Beltramello, A., Bianchetti, A., Binetti, G., Bordiga, G., DeCarli, C., Laakso, M.P., Soininen, H., Testa, C., Zanetti, O., Trabucchi, M., 2002b. Radial Width of the Temporal Horn: A Sensitive Measure in Alzheimer's Disease. Am J Neuroradiol 23, 35-47.

90. Frisoni, G. B., Laakso, M.P., Beltramello, A., Geroldi, C., Bianchetti, A., Soininen, H., Trabucchi, M., 1999. Hippocampal and entorhinal cortex atrophy in frontotemporal dementia and Alzheimer's disease. Neurology 52, 91-100.

91. Friston, K. J., Ashburner, J., Frith, C.D., Poline, J.B., Heather, J.D., Frackowiak, R.S.J., 1995a. Spatial Registration and Normalization of Images. Human Brain Mapping 2, 165-189.

92. Friston, K. J., Holmes, A.P., Worsley, K.J., Poline, J.-P., Frith, C.D., Frackowiak, R.S.J., 1995b. Statistical parametric maps in functional imaging: A general linear approach. Human Brain Mapping 2, 189-210.

93. Frost, C., Kenward, M.G., Fox, N.C., 2004. The analysis of repeated 'direct' measures of change illustrated with an application in longitudinal imaging. Stat.Med 23, 3275-3286.

94. Galanaud, D., Nicoli, F., Le Fur, Y., Guye, M., Ranjeva, J.P., Confort-Gouny, S., Viout, P., Soulier, E., Cozzone, P.J., 2003. Multimodal magnetic resonance imaging of the central nervous system. Biochimie 85, 905-914.

95. Galasko, D., Hansen, L.A., Katzman, R., Wiederholt, W., Masliah, E., et al, 1994. Clinical-Neuropathological correlations in Alzheimer's Disease and related dementias. Archives of Neurology 51, 888-895.

96. Galton, C. J., Patterson, K., Graham, K.S., Lambon-Ralph, M.A., Williams, G., Antoun, N., Sahakian, B.J., Hodges, J.R., 2001. Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. Neurology 57, 216-225.

97. Galton, C. J., Patterson, K., Xuereb, J.H., Hodges, J.R., 2000. Atypical and typical presentations of Alzheimer's disease: a clinical, neuropsychological, neuroimaging and pathological study of 13 cases. Brain 123, 484-498.

98. Geroldi, C., Laakso, M.P., DeCarli, C., Beltramello, A., Bianchetti, A., Soininen, H., Trabucchi, M., Frisoni, G.B., 2000. Apolipoprotein E genotype and hippocampal asymmetry in Alzheimer's disease: a volumetric MRI study. J Neurol Neurosurg Psychiatry 68, 93-96.

99. Ghanei, A., Soltanian-Zadeh, H., Windham, J.P., 1998. Segmentation of the hippocampus from brain MRI using deformable contours. Comput.Med Imaging Graph. 22, 203-216.

100. Giannakopoulos, P., Herrmann, F.R., Bussiere, T., Bouras, C., Kovari, E., Perl, D.P., Morrison, J.H., Gold, G., Hof, P.R., 2003. Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. Neurology 60, 1495-1500.

101. Gilman, S., Koller, M., Black, R.S., Jenkins, L., Griffith, S.G., Fox, N.C., Eisner, L., Kirby, L., Rovira, M.B., Forette, F., Orgogozo, J.M., 2005. Clinical effects of Abeta immunization (AN1792) in patients with AD in an interrupted trial. Neurology 64, 1553-1562.

102. Golebiowski, M., Barcikowska, M., Pfeffer, A., 1999. Magnetic resonance imaging-based hippocampal volumetry in patients with dementia of the Alzheimer type. Dementia And Geriatric Cognitive Disorders 10, 284-288.

103. Gomez-Isla, T., Price, J.L., McKeel-DW, J., Morris, J.C., Growdon, J.H., Hyman, B.T., 1996. Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. J.Neurosci. 16, 4491-4500.

104. Good, C. D., Johnsrude, I.S., Ashburner, J., Henson, R.N.A., Friston, K.J., Frakowiak, R.S.J., 2001. A Voxel-Based Morphometric Study of Ageing in 465 Normal Adult Human Brains. Neuroimage 14, 21-36.

105. Gosche, K. M., Mortimer, J.A., Smith, C.D., Markesbery, W.R., Snowdon, D.A., 2001. An automated technique for measuring hippocampal volumes from MR imaging studies. AJNR Am J Neuroradiol. 22, 1686-1689.

106. Gosche, K. M., Mortimer, J.A., Smith, C.D., Markesbery, W.R., Snowdon, D.A., 2002. Hippocampal volume as an index of Alzheimer neuropathology: findings from the Nun Study. Neurology 58, 1476-1482.

107. Graham, A., Davies, R., Xuereb, J., Halliday, G., Kril, J., Creasey, H., Graham, K., Hodges, J., 2005. Pathologically proven frontotemporal dementia presenting with severe amnesia. Brain 128, 597-605.

108. Green, M. S., Kaye, J.A., Ball, M.J., 2000. The Oregon brain aging study: neuropathology accompanying healthy aging in the oldest old. Neurology 54, 105-113.

109. Groenewegen, H. J., Uylings, H.B., 2000. The prefrontal cortex and the integration of sensory, limbic and autonomic information. Prog.Brain Res. 126, 3-28.

110. Gunter, J. L., Shiung, M.M., Manduca, A., Jack, C.R., Jr., 2003. Methodological considerations for measuring rates of brain atrophy. J Magn Reson. Imaging 18, 16-24.

111. Gur, R. C., Mozley, P.D., Resnick, S.M., Gottlieb, G.L., Kohn, M., Zimmerman, R., Herman, G., Atlas, S., Grossman, R., Berretta, D., Erwin, R., Gur, R.E., 1991. Gender differences in age effect on brain atrophy measured by magnetic resonance imaging. Proceedings of the National Academy of Sciences of the United States of America 88, 2845-2849.

112. Gustafson, L., Brun, A., Passant, U., 1992. Frontal lobe degeneration of non-Alzheimer type. [Review]. Baillieres Clinical Neurology 1, 559-582.

113. Haller, J. W., Banerjee, A., Christensen, G.E., Gado, M., Joshi, S.C., Miller, M.I., Sheline, Y.I., Vannier, M.W., Csernansky, J.G., 1997. Three-dimensional hippocampal MR morphometry with high-dimensional transformation of a neuroanatomical atlas. Radiology 202, 504-510.

114. Haller, J. W., Christensen, G.E., Joshi, S.C., Newcomer, J.W., Miller, M.I., Csernansky, J.G., Vannier, M.W., 1996. Hippocampal MR Imaging Morphometry by Means of General Pattern Matching. Radiol 199, 787-791.

115. Hammers, A., Allom, R., Koepp, M.J., Free, S.L., Myers, R., Lemieux, L., Mitchell, T.N., Brooks, D.J., Duncan, J.S., 2003. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. Hum.Brain Mapp. 19, 224-247.

116. Hammers, A., Koepp, M.J., Free, S.L., Brett, M., Richardson, M.P., Labbe, C., Cunningham, V.J., Brooks, D.J., Duncan, J., 2002. Implementation and

application of a brain template for multiple volumes of interest. Hum.Brain Mapp. 15, 165-174.

117. Hampel, H., Burger, K., Pruessner, J.C., Zinkowski, R., DeBernardis, J., Kerkman, D., Leinsinger, G., Evans, A.C., Davies, P., Moller, H.J., Teipel, S.J., 2005. Correlation of cerebrospinal fluid levels of tau protein phosphorylated at threonine 231 with rates of hippocampal atrophy in Alzheimer disease. Archives of Neurology 62, 770-773.

118. Hampel, H., Teipel, S.J., Alexander, G.E., Horwitz, B., Teichberg, D., Schapiro, M.B., Rapoport, S.I., 1998. Corpus callosum atrophy is a possible indicator of region- and cell type-specific neuronal degeneration in Alzheimer disease: A magnetic resonance imaging analysis. Archives of Neurology 55, 193-198.

119. Hardy, J., 1997. Amyloid, the presenilins and Alzheimer's disease. Trends in Neurosciences 20, 154-159.

120. Hashimoto, M., Kazui, H., Matsumoto, K., Nakano, Y., Yasuda, M., Mori, E., 2005. Does donepezil treatment slow the progression of hippocampal atrophy in patients with Alzheimer's disease? American Journal of Psychiatry 162, 676-682.

121. Hawkes, D.J. 2001. "Registration Methodology: Introduction." In *Medical Image Registration*. CRC Press. Boca Raton. 11-38.

122. He, J., Christensen, G.E., 2003. Large deformation inverse consistent elastic image registration. Inf Process Med Imaging 438-449.

123. Hill, D.L.G. and Batchelor, P. 2001. "Registration Methodology: Concepts and Algorithms." In *Medical Image Registration*. CRC Press. Boca Raton. 39-70.

124. Hoffman, J. M., Welsh-Bohmer, K.A., Hanson, M., Crain, B., Hulette, C., Earl, N., Coleman, R.E., 2000. FDG PET imaging in patients with pathologically verified dementia. J Nucl.Med 41, 1920-1928.

125. Howard, R., Mellers, J., Petty, R., Bonner, D., Menon, R., Almeida, O., Graves, M., Renshaw, C., Levy, R., 1995. Magnetic resonance imaging volumetric measurements of the superior temporal gyrus, hippocampus,

parahippocampal gyrus, frontal and temporal lobes in late paraphrenia. Psychol Med 25, 495-503.

126. Hsu, Y. Y., Schuff, N., Du, A.T., Mark, K., Zhu, X., Hardin, D., Weiner, M.W., 2002. Comparison of Automated and Manual MRI Volumetry of Hippocampus in Normal Aging and Dementia. J Magn Reson Imaging 16, 305-310.

127. Insausti, R., Amaral, D.G., Cowan, W.M., 1987. The entorhinal cortex of the monkey: II. Cortical afferents. J Comp Neurol 264, 356-395.

128. Jack, C. R., Jr., Garwood, M., Wengenack, T.M., Borowski, B., Curran, G.L., Lin, J., Adriany, G., Grohn, O.H., Grimm, R., Poduslo, J.F., 2004. In vivo visualization of Alzheimer's amyloid plaques by magnetic resonance imaging in transgenic mice without a contrast agent. Magn Reson Med 52, 1263-1271.

129. Jack, C. R., Jr., Petersen, R.C., O'Brien, P.C., Tangalos, E.G., 1992. MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease. Neurology 42, 183-188.

130. Jack, C. R., Petersen, R.C., Xu, Y., O'Brien, P.C., Smith, G.E., Ivnik, R.J., Boeve, B.F., Tangalos, E.G., Kokmen, E., 2000. Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. Neurology 55, 484-489.

131. Jack, C. R., Petersen, R.C., Xu, Y., O'Brien, P.C., Smith, G.E., Ivnik, R.J., Tangalos, E.G., Kokmen, E., 1998. Rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease. Neurology 51, 993-999.

132. Jack, C. R., Petersen, R.C., Xu, Y.C., O'Brien, P.C., Smith, G.E., Ivnik, R.J., Boeve, B.F., Waring, S.C., Tangalos, E.G., Kokmen, E., 1999. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. Neurology 52, 1397-1403.

133. Jack, C. R., Jr., Petersen, R.C., Xu, Y.C., O'Brien, P.C., Waring, S.C., Tangalos, E.G., Smith, G.E., Ivnik, R.J., Thibodeau, S.N., Kokmen, E., 1998. Hippocampal atrophy and apolipoprotein E genotype are independently associated with Alzheimer's disease. Annals of Neurology 43, 303-310.

134. Jack, C. R., Petersen, R.C., Xu, Y.C., Waring, S.C., O'Brien, P.C., Tangalos, E.G., Smith, G.E., Ivnik, R.J., Kokmen, E., 1997. Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. Neurology 49, 786-794.

135. Jack, C. R., Jr., Shiung, M.M., Weigand, S.D., O'Brien, P.C., Gunter, J.L., Boeve, B.F., Knopman, D.S., Smith, G.E., Ivnik, R.J., Tangalos, E.G., Petersen, R.C., 2005. Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. Neurology 65, 1227-1231.

136. Jack, C. R., Slomkowski, M., Gracon, S., Hoover, T.M., Felmlee, J.P., Stewart, K., Xu, Y., Shiung, M., O'Brien, P.C., Cha, R., Knopman, D., Petersen, R.C., 2003. MRI as a biomarker of disease progression in a therapeutic trial of milameline for AD. Neurology 60, 253-260.

137. Jack, C. R., Twomey, C.K., Zinsmeister, A.R., Sharbrough, F.W., Petersen, R.C., Cascino, G.D., 1989. Anterior temporal lobes and hippocampal formations: normative volumetric measurements from MR images in young adults. Radiol 172, 549-554.

138. Jack, C. R. J., Shiung, M.M., Gunter, J.L., O'Brien, P.C., Weigand, S.D., Knopman, D.S., Boeve, B.F., Ivnik, R.J., Smith, G.E., Cha, R.H., Tangalos, E.G., Petersen, R.C., 2004. Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. Neurology 62, 591-600.

139. Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Medical Image Analysis 5, 143-156.

140. Jobst, K. A., Smith, A.D., Szatmari, M., Molyneux, A., Esiri, M.E., King, E., Smith, A., Jaskowski, A., McDonald, B., Wald, N., 1992. Detection in life of confirmed Alzheimer's disease using a simple measurement of medial temporal lobe atrophy by computed tomography. Lancet 340, 1179-1183.

141. Johnson, J. K., Head, E., Kim, R., Starr, A., Cotman, C.W., 1999. Clinical and pathological evidence for a frontal variant of Alzheimer disease. Arch.Neurol. 56, 1233-1239.

142. Josephs, K. A., Jones, A.G., Dickson, D.W., 2004. Hippocampal sclerosis and ubiquitin-positive inclusions in dementia lacking distinctive histopathology. Dement Geriatr Cogn Disord 17, 342-345.

143. Juottonen, K., Laakso, M.P., Insausti, R., Lehtovirta, M., Pitkanen, A., Partanen, K., Soininen, H., 1998. Volumes of the entorhinal and perirhinal cortices in Alzheimer's disease. Neurobiol.Aging 19, 15-22.

144. Juottonen, K., Laakso, M.P., Partanen, K., Soininen, H., 1999. Comparative MR analysis of the entorhinal cortex and hippocampus in diagnosing Alzheimer disease. Am J Neuroradiol 20, 139-144.

145. Karas, G. B., Burton, E.J., Rombouts, S.A., van Schijndel, R.A., O'Brien, J.T., Scheltens, P., McKeith, I.G., Williams, D., Ballard, C., Barkhof, F., 2003. A comprehensive study of gray matter loss in patients with Alzheimer's disease using optimized voxel-based morphometry. Neuroimage 18, 895-907.

146. Kaye, J. A., Moore, M.M., Dame, A., Quinn, J., Camicioli, R., Howieson, D., Corbridge, E., Care, B., Nesbit, G., Sexton, G., 2005. Asynchronous regional brain volume losses in presymptomatic to moderate AD. J Alzheimers.Dis 8, 51-56.

147. Kaye, J. A., Swihart, T., Howieson, D., Dame, A., Moore, M.M., Karnos, T., Camicioli, R., Ball, M., Oken, B., Sexton, G., 1997. Volume loss of the hippocampus and temporal lobe in healthy elderly persons destined to develop dementia. Neurology 48, 1297-1304.

148. Kelly, C. A., Harvey, R.J., Cayton, H., 1997. Drug treatments for Alzheimer's disease. British Medical Journal 314, 693-694.

149. Kenward, M. G., Roger, J.H., 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53, 983-997.

150. Kikinis, R., Shenton, M.E., Iosifescu, D.V., McCarley, R.W., Saiviroonporn, P., Hokama, H.H., Robatino, A., Metcalf, D., Wible, C.G., Portas, C.M., Donnino, R.M., Jolesz, F.A., 1996. A digital brain atlas for surgical planning, Model-Driven Segmentation, and Teaching. IEEE Transactions On Visualization and Computer Graphics 2, 232-241.

151. Killiany, R. J., Gomez-Isla, T., Moss, M., Kikinis, R., Sandor, T., Jolesz, F., Tanzi, R., Jones, K., Hyman, B.T., Albert, M.S., 2000. Use of structural magnetic resonance imaging to predict who will get Alzheimer's disease. Annals of Neurology 47, 430-439.

152. Killiany, R. J., Hyman, N., Gomez-Isla, T., 2002. MRI measures of entorhinal cortex vs hippocampus in preclinical AD. Neurology 58, 1188-1196.

153. Killiany, R. J., Moss, M.B., Albert, M.S., Sandor, T., Tieman, J., Jolesz, F., 1993. Temporal lobe regions on magnetic resonance imaging identify patients with early Alzheimer's disease. Archives of Neurology 50, 949-954.

154. Kinsella, K., Phillips, D.R., 2005. Global Aging: The Challenge of Success. Population Bulletin 60.

155. Kizu, O., Yamada, K., Ito, H., Nishimura, T., 2004. Posterior cingulate metabolic changes in frontotemporal lobar degeneration detected by magnetic resonance spectroscopy. Neuroradiology 46, 277-281.

156. Klunk, W. E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D.P., Bergstrom, M., Savitcheva, I., Huang, G.F., Estrada, S., Ausen, B., Debnath, M.L., Barletta, J., Price, J.C., Sandell, J., Lopresti, B.J., Wall, A., Koivisto, P., Antoni, G., Mathis, C.A., Langstrom, B., 2004. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. Annals of Neurology 55, 306-319.

157. Knopman, D. S., Boeve, B.F., Parisi, J.E., Dickson, D.W., Smith, G.E., Ivnik, R.J., Josephs, K.A., Petersen, R.C., 2005. Antemortem diagnosis of frontotemporal lobar degeneration. Annals of Neurology 57, 480-488.

158. Knopman, D. S., DeKosky, S.T., Cummings, J.L., Chui, H., Corey-Bloom, J., Relkin, N., Small, G.W., Miller, B., Stevens, J.C., 2001. Practice parameter: diagnosis of dementia (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 56, 1143-1153.

159. Kobayashi, Y., Amaral, D.G., 2003. Macaque monkey retrosplenial cortex: II. Cortical afferents. J Comp Neurol 466, 48-79.

160. Kochunov, P., Lancaster, J., Thompson, P., Toga, A.W., Brewer, P., Hardies, J., Fox, P., 2002. An optimized individual target brain in the Talairach coordinate system. Neuroimage 17, 922-927.

161. Koeppe, R. A., Gilman, S., Joshi, A., Liu, S., Little, R., Junck, L., Heumann, M., Frey, K.A., Albin, R.L., 2005. 11C-DTBZ and 18F-FDG PET measures in differentiating dementias. J Nucl.Med 46, 936-944.

162. Kohler, S., Black, S.E., Sinden, M., Szekely, C., Kidron, D., Parker, J.L., Foster, J.K., Moscovitch, M., Winocour, G., Szalai, J.P., Bronskill, M.J., 1998. Memory impairments associated with hippocampal versus parahippocampal gyrus atrophy: an MR volumetry study in Alzheimer's disease. Neuropsychologia 36, 901-914.

163. Krasuski, J. S., Alexander, G.E., Horwitz, B., Daly, E.M., Murphy, D.G., Rapoport, S.I., Schapiro, M.B., 1998. Volumes of medial temporal lobe structures in patients with Alzheimer's disease and mild cognitive impairment (and in healthy controls). Biological Psychiatry 43, 60-68.

164. Krishnan, K. R., Charles, H.C., Doraiswamy, P.M., Mintzer, J., Weisler, R., Yu, X., Perdomo, C., Ieni, J.R., Rogers, S., 2003. Randomized, placebo-controlled trial of the effects of donepezil on neuronal markers and hippocampal volumes in Alzheimer's disease. American Journal of Psychiatry 160, 2003-2011.

165. Laakso, M. P., Frisoni, G.B., Kononen, M., Mikkonen, M., Beltramello, A., Geroldi, C., Bianchetti, A., Trabucchi, M., Soininen, H., Aronen, H.J., 2000a. Hippocampus and entorhinal cortex in frontotemporal dementia and Alzheimer's disease: a morphometric MRI study. Biological Psychiatry 47, 1056-1063.

166. Laakso, M. P., Lehtovirta, M., Partanen, K., Riekkinen, P.J., Soininen, H., 2000b. Hippocampus in Alzheimer's disease: a 3-year follow-up MRI study. Biological Psychiatry 47, 557-561.

167. Laakso, M. P., Soininen, H., Partanen, K., Helkala, E.L., Hartikainen, P., Vainio, P., Hallikainen, M., Hanninen, T., Riekkinen, PJ.S., 1995. Volumes of hippocampus, amygdala and frontal lobes in the MRI-based diagnosis of

early Alzheimer's disease: correlation with memory functions. Journal of Neural Transmission 9, 73-86.

168. Lavenex, P., Suzuki, W.A., Amaral, D.G., 2002. Perirhinal and parahippocampal cortices of the macaque monkey: projections to the neocortex. J Comp Neurol 447, 394-420.

169. Lehericy, S., Baulac, M., Chiras, J., Pierot, L., Martin, N., Pillon, B., Deweer, B., Dubois, B., Marsault, C., 1994. Amygdalohippocampal MR volume measurements in the early stages of Alzheimer disease. Am J Neuroradiol 15, 929-937.

170. Lehtovirta, M., Laakso, M.P., Soininen, H., Helisalmi, S., Mannermaa, A., Helkala, E.L., Partanen, K., Ryynanen, M., Vainio, P., Hartikainen, P., Riekkinen, S., 1995. Volumes of hippocampus, amygdala and frontal lobe in Alzheimer patients with different apolipoprotein E genotypes. Neuroscience 67, 65-72.

171. Lewis, E. B., Fox, N.C., 2004. Correction of differential intensity inhomogeneity in longitudinal MR images. Neuroimage 23, 75-83.

172. Likeman, M., Anderson, V.M., Stevens, J.M., Waldman, A.D., Godbolt, A.K., Frost, C., Rossor, M.N., Fox, N.C., 2005. Visual assessment of atrophy on magnetic resonance imaging in the diagnosis of pathologically confirmed young-onset dementias. Archives of Neurology 62, 1410-1415.

173. Ma, C., Wang, G.Z., Braak, H., 1994. Pathological changes of the retrosplenial cortex in senile dementia of Alzheimer type. Chinese Medical Journal - Peking 107, 119-123.

174. Marquis, S., Moore, M.M., Howieson, D.B., Sexton, G., Payami, H., Kaye, J.A., Camicioli, R., 2002. Independent predictors of cognitive decline in healthy elderly persons. Archives of Neurology 59, 601-606.

175. Masdeu, J. C., Zubieta, J.L., Arbizu, J., 2005. Neuroimaging as a marker of the onset and progression of Alzheimer's disease. J Neurol Sci. 236, 55-64.

176. Masliah, E., Hansen, L., Adame, A., Crews, L., Bard, F., Lee, C., Seubert, P., Games, D., Kirby, L., Schenk, D., 2005. Abeta vaccination effects on plaque

pathology in the absence of encephalitis in Alzheimer disease. Neurology 64, 129-131.

177. Matsuda, H., Kitayama, N., Ohnishi, T., Asada, T., Nakano, S., Sakamoto, S., Imabayashi, E., Katoh, A., 2002. Longitudinal evaluation of both morphologic and functional changes in the same individuals with Alzheimer's disease. J Nucl.Med 43, 304-311.

178. Mazziotta, J. C., Pelizzari, C.C., Chen, G.T., Bookstein, F.L., Valentino, D., 1991. Region of interest issues: the relationship between structure and function in the brain. J Cereb.Blood Flow Metab. 11, A51-6.

179. Mazziotta, J. C., Toga, A.W., Evans, A., Fox, P., Lancaster, J., 1995. A probabilistic atlas of the human brain: Theory and rationale for its development. Neuroimage 2, 89-101.

180. McKeith, L. G., Galasko, D., Kosaka, K., Perry, E.K., Dickson, D.W., Hansen, L.A., Salmon, D.P., Lowe, J., Mirra, S.S., Byrne, E.J., Lennox, G., Quinn, N.P., Edwardson, J.A., Ince, P.G., Bergeron, C., Burns, A., Miller, B.L., Lovestone, S., Collerton, D., Jansen, E.N., Ballard, C., de, V.R., Wilcock, G.K., Jellinger, K.A., Perry, R.H., 1996. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. Neurology 47, 1113-1124.

181. McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., Stadlan, E.M., 1984. Clinical diagnosis of Alzheimer's Disease: Report of the NINCDS- ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 34, 939-944.

182. McKhann, G. M., Albert, M.S., Grossman, M., Miller, B., Dickson, D., Trojanowski, J.Q., 2001. Clinical and pathological diagnosis of frontotemporal dementia: report of the Work Group on Frontotemporal Dementia and Pick's Disease. Arch.Neurol. 58, 1803-1809.

183. Mega, M. S., Small, G.W., Xu, M.L., Felix, J., Manese, M., Tran, N.P., Dailey, J.I., Ercoli, L.M., Bookheimer, S.Y., Toga, A.W., 2002. Hippocampal

atrophy in persons with age-associated memory impairment: volumetry within a common space. Psychosom.Med 64, 487-492.

184. Mervaala, E., Fohr, J., Kononen, M., Valkonen, K.M., Vainio, P., Partanen, K., Partanen, J., Tiihonen, J., Viinamaki, H., Karjalainen, A.K., Lehtonen, J., 2000. Quantitative MRI of the hippocampus and amygdala in severe depression. Psychol Med 30, 117-125.

185. Mesulam, M. M., 2001. Primary progressive aphasia. Annals of Neurology 49, 425-432.

186. Minoshima, S., Giordani, B., Berent, S., Frey, K.A., Foster, N.L., Kuhl, D.E., 1997. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. Annals of Neurology 42, 85-94.

187. Mirra, S. S., Hart, M.N., Terry, R.D., 1993. Making the diagnosis of Alzheimer's disease. A primer for practicing pathologists. Arch Pathol Lab Med 117, 132-144.

188. Mirra, S. S., Heyman, A., McKeel, D., Sumi, S.M., Crain, B.J., Brownlee, L.M., Vogel, F.S., Hughes, J.P., VanBelle, G., Berg, L., 1991. The consortium to establish a registry for Alzheimer's disease (CERAD). Standardization of the neuropathologic assessment of Alzheimer's disease. Neurology 41, 479-486.

189. Moffat, S. D., Szekely, C.A., Zonderman, A.B., Kabani, N.J., Resnick, S.M., 2000. Longitudinal change in hippocampal volume as a function of apolipoprotein E genotype. Neurology 55, 134-136.

190. Mori, E., Lee, K., Yasuda, M., Hashimoto, M., Kazui, H., Hirono, N., Matsui, M., 2002. Accelerated hippocampal atrophy in Alzheimer's disease with apolipoprotein E epsilon4 allele. Annals of Neurology 51, 209-214.

191. Morris, J. C., Storandt, M., Miller, J.P., McKeel, D.W., Price, J.L., Rubin, E.H., Berg, L., 2001. Mild cognitive impairment represents early-stage Alzheimer disease. Archives of Neurology 58, 397-405.

192. Mortimer, J. A., Gosche, K.M., Riley, K.P., Markesbery, W.R., Snowdon, D.A., 2004. Delayed recall, hippocampal volume and Alzheimer neuropathology: findings from the Nun Study. Neurology 62, 428-432.

193. Mosconi, L., Sorbi, S., Nacmias, B., De Cristofaro, M.T., Fayyaz, M., Cellini, E., Bagnoli, S., Bracco, L., Herholz, K., Pupi, A., 2003. Brain metabolic differences between sporadic and familial Alzheimer's disease. Neurology 61, 1138-1140.

194. Mosconi, L., Tsui, W.H., De Santi, S., Li, J., Rusinek, H., Convit, A., Li, Y., Boppana, M., de Leon, M.J., 2005. Reduced hippocampal metabolism in MCI and AD: automated FDG-PET image analysis. Neurology 64, 1860-1867.

195. Mueller, E. A., Moore, M.M., Kerr, D.C., Sexton, G., Camicioli, R.M., Howieson, D.B., Quinn, J.F., Kaye, J.A., 1998. Brain volume preserved in healthy elderly through the eleventh decade. Neurology 51, 1555-1562.

196. Mungas, D., Harvey, D., Reed, B.R., Jagust, W.J., DeCarli, C., Beckett, L., Mack, W.J., Kramer, J.H., Weiner, M.W., Schuff, N., Chui, H.C., 2005. Longitudinal volumetric MRI change and rate of cognitive decline. Neurology 65, 565-571.

197. Nagao, M., Sugawara, Y., Ikeda, M., Fukuhara, R., Hokoishi, K., Murase, K., Mochizuki, T., Miki, H., Kikuchi, T., 2004. Heterogeneity of cerebral blood flow in frontotemporal lobar degeneration and Alzheimer's disease. Eur J Nucl.Med Mol.Imaging 31, 162-168.

198. Neary, D., Snowden, J., Mann, D., 2005. Frontotemporal dementia. Lancet Neurol 4, 771-780.

199. Neary, D., Snowden, J.S., Gustafson, L., Passant, U., Stuss, D., Black, S., Freedman, M., Kertesz, A., Robert, P.H., Albert, M., Boone, K., Miller, B.L., Cummings, J., Benson, D.F., 1998. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. Neurology 51, 1546-1554.

200. Newell, K. L., Hyman, B.T., Growdon, J.H., Hedley-Whyte, E.T., 1999. Application of the National Institute on Aging (NIA)-Reagan Institute criteria for the neuropathological diagnosis of Alzheimer disease. J Neuropathol.Exp.Neurol 58, 1147-1155.

201. Nicoll, J. A., Wilkinson, D., Holmes, C., Steart, P., Markham, H., Weller, R.O., 2003. Neuropathology of human Alzheimer disease after immunization with amyloid-beta peptide: a case report. Nat.Med 9, 448-452.

202. O'Brien, J. T., Paling, S., Barber, R., Williams, E.D., Ballard, C., McKeith, I.G., Gholkar, A., Crum, W.R., Rossor, M.N., Fox, N.C., 2001. Progressive brain atrophy on serial MRI in dementia with Lewy bodies, AD, and vascular dementia. Neurology 56, 1386-1388.

203. O'Riordan, S., McMonagle, P., Janssen, J.C., Fox, N.C., Farrell, M., Collinge, J., Rossor, M.N., Hutchinson, M., 2002. Presenilin-1 mutation (E280G), spastic paraparesis, and cranial MRI white-matter abnormalities. Neurology 59, 1108-1110.

204. Ono,M., Kubic, S., and Abernathey, C.D., 1990. *Atlas of the cerebral sulci.* Thieme Medical Publishers.

205. Orgogozo, J. M., Gilman, S., Dartigues, J.F., Laurent, B., Puel, M., Kirby, L.C., Jouanny, P., Dubois, B., Eisner, L., Flitman, S., Michel, B.F., Boada, M., Frank, A., Hock, C., 2003. Subacute meningoencephalitis in a subset of patients with AD after Abeta42 immunization. Neurology 61, 46-54.

206. Ostuni, J. L., Levin, R.L., Frank, J.A., DeCarli, C., 1997. Correspondence of closest gradient voxels--a robust registration algorithm. J Magn Reson Imaging 7, 410-415.

207. Paus, T., Tomaiuolo, F., Otaky, N., MacDonald, D., Petrides, M., Atlas, J., Morris, R., Evans, A.C., 1996. Human cingulate and paracingulate sulci: pattern, variability, asymmetry, and probabilistic map. Cereb.Cortex 6, 207-214.

208. Pennanen, C., Kivipelto, M., Tuomainen, S., Hartikainen, P., Hanninen, T., Laakso, M.P., Hallikainen, M., Vanhanen, M., Nissinen, A., Helkala, E.L., Vainio, P., Vanninen, R., Partanen, K., Soininen, H., 2004. Hippocampus and entorhinal cortex in mild cognitive impairment and early AD. Neurobiol Aging 25, 303-310.

209. Petersen, R. C., Doody, R., Kurz, A., Mohs, R.C., Morris, J.C., Rabins, P.V., Ritchie, K., Rossor, M., Thal, L., Winblad, B., 2001a. Current concepts in mild cognitive impairment. Arch.Neurol. 58, 1985-1992.

210. Petersen, R. C., Smith, G.E., Waring, S.C., Ivnik, R.J., Tangalos, E.G., Kokmen, E., 1999. Mild cognitive impairment: clinical characterization and outcome. Archives of Neurology 56, 303-308.

211. Petersen, R. C., Stevens, J.C., Ganguli, M., Tangalos, E.G., Cummings, J.L., DeKosky, S.T., 2001b. Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review). Neurology 56, 1133-1142.

212. Petrella, J. R., Coleman, R.E., Doraiswamy, P.M., 2003. Neuroimaging and early diagnosis of Alzheimer disease: a look to the future. Radiology 226, 315-336.

213. Pitiot, A., Delingette, H., Thompson, P.M., Ayache, N., 2004. Expert knowledge-guided segmentation system for brain MRI. Neuroimage 23 Suppl 1, S85-S96.

214. Pruessner, L. M., Li, L.M., Serles, W., Pruessner, M., Collins, D.L., Kabani, N., Lupien, S., Evans, A.C., 2000. Volumetry of Hippocampus and Amygdala with High-resolution MRI and Three-dimensional Analysis Software: Minimizing the Discrepancies between Laboratories. Cereb Cortex 10, 433-442.

215. Pucci, E., Belardinelli, N., Regnicolo, L., Nolfe, G., Signorino, M., Salvolini, U., Angeleri, F., 1998. Hippocampus and parahippocampal gyrus linear measurements based on magnetic resonance in Alzheimer's disease. Journal of European Neurology 39, 16-25.

216. Rabiner, L. R., Schafer, R.W., Rader, C.M., 1969. The Chirp z-Transformation algorithm and its application. The Bell System Technical Journal 48, 1249-1292.

217. Ratnavalli, E., Brayne, C., Dawson, K., Hodges, J.R., 2002. The prevalence of frontotemporal dementia. Neurology 58, 1615-1621.

218. Reiman, E. M., Uecker, A., Caselli, R.J., Lewis, S., Bandy, D., de Leon MJ, De Santi S, Convit, A., Osborne, D., Weaver, A., Thibodeau, S.N., 1998. Hippocampal volumes in cognitively normal persons at genetic risk for Alzheimer's disease. Annals of Neurology 44, 288-291.

219. Resnick, S. M., Goldszal, A.F., Davatzikos, C., Golski, S., Kraut, M.A., Metter, E.J., Bryan, R.N., Zonderman, A.B., 2000. One-year age changes in MRI brain volumes in older adults. Cerebral Cortex 10, 464-472.

220. Rohn, T. T., Head, E., Su, J.H., Anderson, A.J., Bahr, B.A., Cotman, C.W., Cribbs, D.H., 2001. Correlation between caspase activation and neurofibrillary tangle formation in Alzheimer's disease. Am J Pathol. 158, 189-198.

221. Roman, G. C., Tatemichi, T.K., Erkinjuntti, T., Cummings, J.L., Masdeu, J.C., Garcia, J.H., Amaducci, L., Orgogozo, J.M., Brun, A., Hofman, A., Moody, D.M., O'Brien, M.D., Yamaguchi, T., Grafman, J., Drayer, B.P., Bennett, D.A., Fisher, M., Ogata, J., Kokmen, E., Bermejo, F., Wolf, P.A., Gorelick, P.B., Bick, K.L., Pajeau, A.K., Bell, M.A., DeCarli, C., Culebras, A., Korczyn, A.D., Bogousslavsky, J., Hartmann, A., Scheinberg, P., 1993. Vascular Dementia: Diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. Neurology 43, 250-260.

222. Rombouts, S. A., Barkhof, F., Witter, M.P., Scheltens, P., 2000. Unbiased whole-brain analysis of gray matter loss in Alzheimer's disease. Neuroscience Letters 285, 231-233.

223. Rosen, H. J., Gorno-Tempini, M.L., Goldman, W.P., Perry, R.J., Schuff, N., Weiner, M.W., Feiwell, R., Kramer, J.H., Miller, B.L., 2002a. Patterns of brain atrophy in frontotemporal dementia and semantic dementia. Neurology 58, 198-208.

224. Rosen, H. J., Hartikainen, K.M., Jagust, W., Kramer, J.H., Reed, B.R., Cummings, J.L., Boone, K., Ellis, W., Miller, C., Miller, B.L., 2002b. Utility of clinical criteria in differentiating frontotemporal lobar degeneration (FTLD) from AD. Neurology 58, 1608-1615.

225. Rosen, H. J., Perry, R.J., Murphy, J., Kramer, J.H., Mychack, P., Schuff, N., Weiner, M., Levenson, R.W., Miller, B.L., 2002c. Emotion comprehension in the temporal variant of frontotemporal dementia. Brain 125, 2286-2295.

226. Rosen, H. J., Wilson, M.R., Schauer, G.F., Allison, S., Gorno-Tempini, M.L., Pace-Savitsky, C., Kramer, J.H., Levenson, R.W., Weiner, M., Miller, B.L., 2006. Neuroanatomical correlates of impaired recognition of emotion in dementia. Neuropsychologia 44, 365-373.

227. Rossor, M. N., 1993. Alzheimer's disease. British Medical Journal 307, 779-782.

228. Rossor, M. N., Fox, N.C., Freeborough, P.A., Harvey, R.J., 1996. Clinical features of sporadic and familial Alzheimer's disease. Neurodegeneration. 5, 393-397.

229. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imaging 18, 712-721.

230. Rusinek, H., de Santi, D., Frid, D., Tsui, W.H., Tarshish, C.Y., Convit, A., de Leon, M.J., 2003. Regional brain atrophy rate predicts future cognitive decline: 6-year longitudinal MR imaging study of normal aging. Radiology. 229, 691-696.

231. Sanz Arigita, E.J., de Vos, K., Pool, C.W., Zilles, K., Uylings, H.B.M., 2003. Individual variability of human anterior cingulate cortex combined with 3D MRI, Abstract from the 9th International Conference on Functional Mapping of the Human Brain.

232. Scahill, R. I., Frost C, Jenkins R, Whitwell J.L., Rossor M.N., Fox N.C., 2003. A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. Arch Neurol. 60, 989-994.

233. Scahill, R. I., Schott, J.M., Stevens, J.M., Rossor, M.N., Fox, N.C., 2002. Mapping the evolution of regional atrophy in Alzheimer's disease: unbiased analysis of fluid-registered serial MRI. Proceedings of the National Academy of Sciences of the United States of America 99, 4703-4707.

234. Scarmeas, N., Habeck, C.G., Zarahn, E., Anderson, K.E., Park, A., Hilton, J., Pelton, G.H., Tabert, M.H., Honig, L.S., Moeller, J.R., Devanand, D.P., Stern, Y., 2004. Covariance PET patterns in early Alzheimer's disease and subjects with cognitive impairment but no dementia: utility in group discrimination and correlations with functional performance. Neuroimage 23, 35-45.

235. Scheltens, P., Fox, N.C., Barkhof, F., DeCarli, C.D., 2002. Structural magnetic resonance imaging in the practical assessment of dementia: beyond exclusion. Lancet Neurology 1, 13-21.

236. Scheltens, P., Erkinjunti, T., Leys, D., Wahlund, L.O., Inzitari, D., del Ser, T., Pasquier, F., Barkhof, F., Mantyla, R., Bowler, J., Wallin, A., Ghika, J., Fazekas, F., Pantoni, L., 1998. White matter changes on CT and MRI: an overview of visual rating scales. European Task Force on Age-Related White Matter Changes. Journal of European Neurology 39, 80-89.

237. Scheltens, P., Leys, D., Barkhof, F., Huglo, D., Weinstein, H.C., Vermersch, P., Kuiper, M., Steinling, M., Wolters, E.C., Valk, J., 1992. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. J Neurol Neurosurg Psychiatry 55, 967-972.

238. Schott, J. M., Fox, N.C., Frost, C., Scahill, R.I., Janssen, J.C., Chan, D., Jenkins, R., Rossor, M.N., 2003. Assessing the onset of structural change in familial Alzheimer's disease. Ann.Neurol. 53, 181-188.

239. Schott, J. M., Price, S.L., Frost, C., Whitwell, J.L., Rossor, M.N., Fox, N.C., 2005. Measuring atrophy in Alzheimer disease - A serial MRI study over 6 and 12 months. Neurology 65, 119-124.

240. Sheline, Y. I., Gado, M.H., Price, J.L., 1998. Amygdala core nuclei volumes are decreased in recurrent major depression. Neuroreport 9, 2023-2028.

241. Shen, D., Moffat, S., Resnick, S.M., Davatzikos, C., 2002. Measuring size and shape of the hippocampus in MR images using a deformable shape model. Neuroimage 15, 422-434.

242. Silbert, L. C., Quinn, J.F., Moore, M.M., Corbridge, E., Ball, M.J., Murdoch, G., Sexton, G., Kaye, J.A., 2003. Changes in premorbid brain volume predict Alzheimer's disease pathology. Neurology 61, 487-492.

243. Silverman, D. H., 2004. Brain 18F-FDG PET in the diagnosis of neurodegenerative dementias: comparison with perfusion SPECT and with clinical evaluations lacking nuclear imaging. J Nucl.Med 45, 594-607.

244. Sled, J. G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans Med Imaging 17, 87-97.

245. Small, G. W., Ercoli, L.M., Silverman, D.H., Huang, S.-C., Komo, S., Bookheimer, S.Y., Lavretsky, H., Miller, K., Siddarth, P., Rasgon, N.L., Mazziotta, J.C., Saxena, S., Wu, H.M., Mega, M.S., Cummings, J.L., Saunders, A.M., Pericak-Vance, M.A., Roses, A.D., Barrio, J.R., Phelps, M.E., 2000. Cerebral metabolic and cognitive decline in persons at genetic risk for Alzheimer's disease. Proceedings of the National Academy of Sciences of the United States of America 97, 6037-6042.

246. Small, G. W., Rabins, P.V., Barry, P.P., Buckholtz, N.S., DeKosky, S.T., Ferris, S.H., Finkel, S.I., Gwyther, L.P., Khachaturian, Z.S., Lebowitz, B.D., McRae, T.D., Morris, J.C., Oakley, F., Schneider, L.S., Streim, J.E., Sunderland, T., Teri, L.A., Tune, L.E., 1997. Diagnosis and treatment of Alzheimer disease and related disorders. Consensus statement of the American Association for Geriatric Psychiatry, the Alzheimer's Association, and the American Geriatrics Society. JAMA 278, 1363-1371.

247. Smith, S. M., De Stefano, N., Jenkinson, M., Matthews, P.M., 2001. Normalized accurate measurement of longitudinal brain change. J Comput.Assist.Tomogr. 25, 466-475.

248. Soininen, H., Partanen, K., Pitkanen, A., Hallikainen, M., Hanninen, T., Helisalmi, S., Mannermaa, A., Ryynanen, M., Koivisto, K., Riekkinen, S., 1995. Decreased hippocampal volume asymmetry on MRIs in nondemented elderly subjects carrying the apolipoprotein E epsilon 4 allele. Neurology 45, 391-392.

249. Soininen, H. S., Partanen, K., Pitkanen, A., Vainio, P., Hanninen, T., Hallikainen, M., Koivisto, K., Riekkinen, P.J., 1994. Volumetric MRI analysis of the amygdala and the hippocampus in subjects with age-associated memory impairment: Correlation to visual and verbal memory. Neurology 44, 1660-1668.

250. Sowell, E. R., Thompson, P.M., Toga, A.W., 2004. Mapping changes in the human cortex throughout the span of life. Neuroscientist. 10, 372-392.

251. Stoub, T. R., deToledo-Morrell, L., Stebbins, G.T., Leurgans, S., Bennett, D.A., Shah, R.C., 2006. Hippocampal disconnection contributes to memory dysfunction in individuals at risk for Alzheimer's disease. Proc Natl.Acad.Sci.U.S.A 103, 10041-10045.

252. Studholme, C., Cardenas-Nicolson, V., Weiner, M., 2001. Building Whole Brain Maps of Atrophy Rate from Multi-Subject Longitudinal Studies Using Free Form Deformations. Proceedings of the International Society for Magnetic Resonance in Medicine 9, 818.

253. Suzuki, W. A., Amaral, D.G., 1994. Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. J Comp Neurol 350, 497-533.

254. Talairach, J. and Tournoux, P. 1988. *Co-planar Stereotaxic Atlas of the Human Brain*. Thieme Verlag. Stuttgart.

255. Thacker, N. A., Jackson, A., Moriarty, D., Vokurka, E.A., 1999. Improved Quality of Re-sliced MR Images Using Re-normalized Sinc Interpolation. Journal of Magnetic Resonance Imaging 10, 582-588.

256. The Lund and Manchester Groups, 1994. Clinical and neuropathological criteria for frontotemporal dementia. The Lund and Manchester Groups. Journal Of Neurology Neurosurgery And Psychiatry 57, 416-418.

257. Thevenaz, P., Ruttimann, U.E., Unser, M., 1998. A Pyramid Approach to Subpixel Registration Based on Intensity. IEEE Trans Image Proc 7, 27-41.

258. Thompson, P. M., Hayashi, K.M., de Zubicaray, G., Janke, A.L., Rose, S.E., Semple, J., Herman, D., Hong, M.S., Dittmer, S.S., Doddrell, D.M., Toga,

A.W., 2003. Dynamics of gray matter loss in Alzheimer's disease. J Neurosci. 23, 994-1005.

259. Thompson, P. M., Hayashi, K.M., de Zubicaray, G.I., Janke, A.L., Rose, S.E., Semple, J., Hong, M.S., Herman, D.H., Gravano, D., Doddrell, D.M., Toga, A.W., 2004a. Mapping hippocampal and ventricular change in Alzheimer disease. Neuroimage 22, 1754-1766.

260. Thompson, P. M., Hayashi, K.M., Sowell, E.R., Gogtay, N., Giedd, J.N., Rapoport, J.L., de Zubicaray, G.I., Janke, A.L., Rose, S.E., Semple, J., Doddrell, D.M., Wang, Y., van Erp, T.G., Cannon, T.D., Toga, A.W., 2004b. Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia. Neuroimage 23 Suppl 1, S2-18.

261. Thompson, P. M., Mega, M.S., Woods, R.P., Zoumalan, C.I., Lindshield, C.J., Blanton, R.E., Moussai, J., Holmes, C.J., Cummings, J.L., Toga, A.W., 2001. Cortical change in Alzheimer's disease detected with a disease-specific population-based brain atlas. Cereb.Cortex 11, 1-16.

262. Thompson, P. M., Moussai, J., Zohoori, S., Goldkorn, A., Khan, A.A., Mega, M.S., Small, G.W., Cummings, J.L., Toga, A.W., 1998. Cortical Variability and Asymmetry in Normal Aging and Alzheimer's Disease. Cereb. Cortex 8, 492-509.

263. Thompson, S. A., Graham, K.S., Williams, G., Patterson, K., Kapur, N., Hodges, J.R., 2004c. Dissociating person-specific from general semantic knowledge: roles of the left and right temporal lobes. Neuropsychologia 42, 359-370.

264. Tisserand, D. J., Pruessner, J.C., Sanz Arigita, E.J., van Boxtel, M.P., Evans, A.C., Jolles, J., Uylings, H.B.M., 2002. Regional frontal cortical volumes decrease differentially in aging: an MRI study to compare volumetric approaches and voxel-based morphometry. Neuroimage 17, 657-669.

265. Tisserand, D. J., van Boxtel, M.P., Pruessner, J.C., Hofman, P., Evans, A.C., Jolles, J., 2004. A voxel-based morphometric study to determine individual differences in gray matter density associated with age and cognitive change over time. Cereb.Cortex 14, 966-973.

266. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273-289.

267. Uylings, H. B., de Brabander, J.M., 2002. Neuronal changes in normal human aging and Alzheimer's disease. Brain Cogn 49, 268-276.

268. Uylings, H. B., Rajkowska, G., Sanz-Arigita, E., Amunts, K., Zilles, K., 2005. Consequences of large interindividual variability for human brain atlases: converging macroscopical imaging and microscopical neuroanatomy. Anat.Embryol.(Berl) 210, 423-431.

269. Uylings, H. B., Sanz, A.E., de Vos, K., Smeets, W.J., Pool, C.W., Amunts, K., Rajkowska, G., Zilles, K., 2000. The importance of a human 3D database and atlas for studies of prefrontal and thalamic functions. Prog.Brain Res. 126, 357-368.

270. Valla, J., Berndt, J.D., Gonzalez-Lima, F., 2001. Energy hypometabolism in posterior cingulate cortex of Alzheimer's patients: superficial laminar cytochrome oxidase associated with disease duration. J.Neurosci. 21, 4923-4930.

271. van de Pol, L. A., Hensel, A., van der Flier, W.M., Visser, P.J., Pijnenburg, Y.A., Barkhof, F., Gertz, H.J., Scheltens, P., 2006. Hippocampal atrophy on MRI in frontotemporal lobar degeneration and Alzheimer's disease. J Neurol Neurosurg Psychiatry 77, 439-442.

272. Vargha-Khadem, F., Gadian, D.G., Watkins, K.E., Connelly, A., Van, P.W., Mishkin, M., 1997. Differential effects of early hippocampal pathology on episodic and semantic memory. Science 277, 376-380.

273. Vogt, B. A., Berger, G.R., Derbyshire, S.W., 2003. Structural and functional dichotomy of human midcingulate cortex. Eur J Neurosci 18, 3134-3144.

274. Vogt, B.A., Hof, P.R., Vogt, L.J., 2004. "Cingulate gyrus." In G.Paxinos and J.K.Mai, editors, *The human nervous system*. Academic Press. San Diego.

275. Vogt, B. A., Nimchinsky, E.A., Vogt, L.J., Hof, P.R., 1995. Human cingulate cortex: surface features, flat maps, and cytoarchitecture. J Comp Neurol 359, 490-506.

276. Vogt, B. A., Rosene, D.L., Pandya, D.N., 1979. Thalamic and cortical afferents differentiate anterior from posterior cingulate cortex in the monkey. Science 204, 205-207.

277. Vogt, B. A., Van Hoesen, G.W., Vogt, L.J., 1990. Laminar distribution of neuron degeneration in posterior cingulate cortex in Alzheimer's disease. Acta Neuropathol (Berl) 80, 581-589.

278. Vogt, B. A., Vogt, L., 2003. Cytology of human dorsal midcingulate and supplementary motor cortices. J Chem.Neuroanat. 26, 301-309.

279. Vogt, B. A., Vogt, L.J., Perl, D.P., Hof, P.R., 2001. Cytology of human caudomedial cingulate, retrosplenial, and caudal parahippocampal cortices. J Comp Neurol 438, 353-376.

280. Vogt, B. A., Vogt, L.J., Vrana, K.E., Gioia, L., Meadows, R.S., Challa, V.R., Hof, P.R., Van Hoesen, G.W., 1998. Multivariate analysis of laminar patterns of neurodegeneration in posterior cingulate cortex in Alzheimer's disease. Exp.Neurol 153, 8-22.

281. Von Gunten, A., Fox, N.C., Cipolotti, L., Ron, M.A., 2000. A volumetric study of hippocampus and amygdala in depressed patients with subjective memory problems. J.Neuropsychiatry Clin.Neurosci. 12, 493-498.

282. Von Gunten, A., Kovari, E., Rivara, C.B., Bouras, C., Hof, P.R., Giannakopoulos, P., 2005. Stereologic analysis of hippocampal Alzheimer's disease pathology in the oldest-old: evidence for sparing of the entorhinal cortex and CA1 field. Exp.Neurol 193, 198-206.

283. Wahlund, L. O., Julin, P., Johansson, S.E., Scheltens, P., 2000. Visual rating and volumetry of the medial temporal lobe on magnetic resonance imaging in dementia: A comparative study. J Neurol Neurosurg Psychiatry 69, 630-635.

284. Waldemar, G., Dubois, B., Emre, M., Scheltens, P., Tariska, P., Rossor, M., 2000. Diagnosis and management of Alzheimer's disease and other disorders

associated with dementia. The role of neurologists in Europe. European Federation of Neurological Societies. Eur.J.Neurol. 7, 133-144.

285. Wang, D., Chalk, J.B., Rose, S.E., de Zubicaray, G., Cowin, G., Galloway, G.J., Barnes, D., Spooner, D., Doddrell, D.M., Semple, J., 2002. MR image-based measurement of rates of change in volumes of brain structures. Part II: application to a study of Alzheimer's disease and normal aging. Magn. Reson.Imaging 20, 41-48.

286. Wang, L., Joshi, S.C., Miller, M.I., Csernansky, J.G., 2001. Statistical Analysis of Hippocampal Asymmetry in Schizophrenia. Neuroimage 14, 531-545.

287. Wang, L., Swank, J.S., Glick, I.E., Gado, M.H., Miller, M.I., Morris, J.C., Csernansky, J.G., 2003. Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. Neuroimage 20, 667-682.

288. Watkins, K. E., Paus, T., Lerch, J.P., Zijdenbos, A., Collins, D.L., Neelin, P., Taylor, J., Worsley, K.J., Evans, A.C., 2001. Structural asymmetries in the human brain: a voxel-based statistical analysis of 142 MRI scans. Cereb.Cortex 11, 868-877.

289. Watson, C., Andermann, F., Gloor, P., Jones-Gotman, M., Peters, T., Evans, A., Olivier, A., Melanson, D., Leroux, G., 1992. Anatomic basis of amygdaloid and hippocampal volume measurement by magnetic resonance imaging. Neurology 42, 1743-1750.

290. Webb, J., Guimond, A., Eldridge, P., Chadwick, D., Meunier, J., Thirion, J.P., Roberts, N., 1999. Automatic detection of hippocampal atrophy on magnetic resonance images. Magn Reson.Imaging 17, 1149-1161.

291. Weinberger, D. R., Luchins, D.J., Morihisa, J., Wyatt, R.J., 1981. Asymmetrical Volumes of the Right and Left Frontal and Occipital Regions of the Human Brain. Ann.Neurol 11, 97-100.

292. Whitwell, J. L., Anderson, V.M., Scahill, R.I., Rossor, M.N., Fox, N.C., 2004a. Longitudinal patterns of regional change on volumetric MRI in frontotemporal lobar degeneration. Dement Geriatr Cogn Disord 17, 307-310.

293. Whitwell, J. L., Crum, W.R., Watt, H.C., Fox, N.C., 2001. Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. AJNR Am.J.Neuroradiol. 22, 1483-1489.

294. Whitwell, J. L., Josephs, K.A., Rossor, M.N., Stevens, J.M., Revesz, T., Holton, J.L., al Sarraj, S., Godbolt, A.K., Fox, N.C., Warren, J.D., 2005a. Magnetic resonance imaging signatures of tissue pathology in frontotemporal dementia. Archives of Neurology 62, 1402-1408.

295. Whitwell, J. L., Sampson, E.L., Watt, H.C., Harvey, R.J., Rossor, M.N., Fox, N.C., 2005b. A volumetric magnetic resonance imaging study of the amygdala in frontotemporal lobar degeneration and Alzheimer's disease. Dement Geriatr Cogn Disord 20, 238-244.

296. Whitwell, J. L., Schott, J.M., Lewis, E.B., MacManus, D.G., Fox, N.C., 2004b. Using nine degrees-of-freedom registration to correct for changes in voxel size in serial MRI studies. Magn Reson Imaging 22, 993-999.

297. Wolfson, L., Wei, X., Hall, C.B., Panzer, V., Wakefield, D., Benson, R.R., Schmidt, J.A., Warfield, S.K., Guttmann, C.R., 2005. Accrual of MRI white matter abnormalities in elderly with normal and impaired mobility. J Neurol Sci. 232, 23-27.

298. Woods, R. P., Grafton, S.T., Holmes, C.J., Cherry, S.R., Mazziotta, J.C., 1998. Automated image registration: I. General methods and intrasubject, intramodality validation. J Comput.Assist.Tomogr. 22, 139-152.

299. Xu, Y., Jack, C.R., Jr., O'Brien, P.C., Kokmen, E., Smith, G.E., Ivnik, R.J., Boeve, B.F., Tangalos, R.G., Petersen, R.C., 2000. Usefulness of MRI measures of entorhinal cortex versus hippocampus in AD. Neurology 54, 1760-1767.

300. Zhilkin, P., Alexander, M.E., 2000. 3D image registration using a fast noniterative algorithm. Magn Reson Imaging 18, 1143-1150.

301. Zilles, K. 2004. "Architecture of the human cerebral cortex. Regional and laminar organization." In *The human nervous system.* Academic Press. San Diego. 997-1055.